GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 24, 2005, 23:19:44 ; Search time 164 Seconds
(without alignments)
37.733 Million cell updates/sec

Title: US-09-865-281A-1
Perfect score: 91
Sequence: 1 KNRWEDPGKQLYNVEA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	100.0	16	4	AAB92360	Aab92360 Miscellan
2	91	100.0	16	6	ABP58217	Abp58217 Immunosti
3	91	100.0	16	8	ADS17594	Ads17594 Peptide d
4	91	100.0	63	5	AAB71451	Aab71451 Human C3
5	91	100.0	294	5	AAU74858	Aau74858 Complemen
6	91	100.0	294	5	AAU74866	Aau74866 Complemen
7	91	100.0	294	5	AAU74869	Aau74869 Complemen
8	91	100.0	294	5	AAU74855	Aau74855 Complemen
9	91	100.0	294	5	AAU74862	Aau74862 Complemen
10	91	100.0	294	5	AAU74859	Aau74859 Complemen
11	91	100.0	294	5	AAU74872	Aau74872 Complemen
12	91	100.0	294	5	AAU74873	Aau74873 Complemen
13	91	100.0	294	5	AAU74863	Aau74863 Complemen
14	91	100.0	294	5	AAU74856	Aau74856 Complemen
15	91	100.0	294	5	AAU74880	Aau74880 Complemen
16	91	100.0	294	5	AAU74860	Aau74860 Complemen
17	91	100.0	294	5	AAU74854	Aau74854 Complemen
18	91	100.0	294	5	AAU74865	Aau74865 Complemen
19	91	100.0	294	5	AAU74867	Aau74867 Complemen
20	91	100.0	294	5	AAU74861	Aau74861 Complemen
21	91	100.0	294	5	AAU74871	Aau74871 Complemen
22	91	100.0	294	5	AAU74868	Aau74868 Complemen
23	91	100.0	294	5	AAU74874	Aau74874 Complemen
24	91	100.0	294	5	AAU74878	Aau74878 Complemen
25	91	100.0	294	5	AAU74879	Aau74879 Complemen
26	91	100.0	294	5	AAU74857	Aau74857 Complemen
27	91	100.0	294	5	AAU74864	Aau74864 Complemen
28	91	100.0	294	5	AAU74870	Aau74870 Complemen
29	91	100.0	294	5	AAU74875	Aau74875 Complemen
30	91	100.0	294	5	AAU74876	Aau74876 Complemen
31	91	100.0	294	5	AAU74877	Aau74877 Complemen
32	91	100.0	294	5	AAU74881	Aau74881 Complemen
33	91	100.0	310	8	ADI05803	Adi05803 Human com
34	91	100.0	310	8	ADI05805	Adi05805 Human C3d
35	91	100.0	310	8	ADI05804	Adi05804 Human C3d
36	91	100.0	349	2	AAR10900	Aar10900 Human pho
37	91	100.0	349	2	AAR21776	Aar21776 Phospholi
38	91	100.0	349	2	AAR51949	Aar51949 Phospholi
39	91	100.0	370	8	ADK72548	Adk72548 Fusion pr
40	91	100.0	383	8	ADK72551	Adk72551 Fusion pr
41	91	100.0	387	8	ADK72549	Adk72549 Fusion pr
42	91	100.0	388	8	ADK72550	Adk72550 Fusion pr
43	91	100.0	705	7	ADD93520	Add93520 Novel NOV
44	91	100.0	1255	6	ABR63374	Abr63374 Human Alz
45	91	100.0	1288	8	ADQ39663	Adq39663 Human myo

ALIGNMENTS

RESULT 1
AAB92360
ID AAB92360 standard; peptide; 16 AA.
XX
AC AAB92360;
XX
DT 22-JUN-2001 (first entry)
XX
DE Miscellaneous peptide SEQ ID NO:1536.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US013576.
XX
PR 17-MAY-1999; 99US-0134406P.
PR 10-SEP-1999; 99US-0153406P.
PR 15-OCT-1999; 99US-0159783P.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity.
XX
PS Disclosure; Page 707; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity in
CC vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent

CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 91; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 1 KNRWEDPGKQLYNVEA 16

RESULT 2
ABP58217
ID ABP58217 standard; peptide; 16 AA.
XX
AC ABP58217;
XX
DT 21-MAR-2003 (first entry)
XX
DE Immunostimulant C3d peptide.
XX
KW Immunostimulant; C3d; human; fusion protein; tumour; vaccine; adjuvant.
XX
OS Homo sapiens.
XX
PN WO200297041-A2.
XX
PD 05-DEC-2002.
XX
PF 29-MAY-2002; 2002WO-US016651.
XX
PR 29-MAY-2001; 2001US-00865281.
XX
PA (IMMP-) IMMPHERON INC.
PA (INNE-) INNEXUS CORP.
XX
PI Kohler H, Morgan C;
XX
DR WPI; 2003-140458/13.
XX
PT Novel fusion protein for use as molecular adjuvant, has an antibody and a
PT peptide with immunostimulatory, membrane transport or homophilic
PT activities, connected to the antibody by peptide bonds.
XX
PS Example 1; Page 14; 39pp; English.
XX
CC The present invention provides a fusion protein made up of an antibody
CC and a peptide having e.g. immunostimulant, membrane transport or
CC homophilic activity. The peptide is located at a site in the antibody
CC such that it does not compromise the antigen recognition of the antibody.
CC In order to enhance its activity, the peptide may be flanked by loop-
CC forming or conformation-conferring sequences. The present sequence is an
CC example of a suitable imunostimulatory peptide for use as a fusion
CC protein component. The peptide is derived from human C3d amino acids 1217
CC -1232. In examples from the invention, the C3d peptide was affinity cross
CC -linked to tumour anti-idiotype and tumour idiotype vaccine antibodies,
CC significantly enhancing the immune response to the tumour and protecting
CC against tumour challenge. The vaccination protocol did not include any
CC adjuvant, such as Freund's adjuvant or keyhole limpet haemocyanin
CC conjugation, both of which are not permissible by the FDA for human use
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 91; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 1 KNRWEDPGKQLYNVEA 16

RESULT 3
ADS17594
ID ADS17594 standard; peptide; 16 AA.
XX
AC ADS17594;
XX
DT 02-DEC-2004 (first entry)
XX
DE Peptide derived from the C3d peptide and affinity linked to 3H1 antibody.
XX
KW immunostimulatory; membrane transport; homophilic; signaling protein;
KW caspase; kinase; phosphatase; viral protein; tumour antigen;
KW nuclear protein; nucleolar protein; DNA synthesis; cytoskeletal protein;
KW cell proliferation; cytostasis; membrane transporter peptide;
KW Kaposi fibroblast factor; TAT peptide; HIV-1; antennapedia homeodomain;
KW herpes virus protein VP22; transportan peptide; Alzheimer's disease;
KW Huntington's disease; Parkinson's disease; C3d; 3H1; monoclonal antibody;
KW anti-idiotypic antibody; carcino-embryonic antigen; CEA;
KW anti-idiotype vaccine; antibody.
XX
OS Synthetic.
XX
PN WO2004078146-A2.
XX
PD 16-SEP-2004.
XX
PF 05-MAR-2004; 2004WO-US006911.
XX
PR 05-MAR-2003; 2003US-0451980P.
XX
PA (INNE-) INNEXUS BIOTECHNOLOGY INC.
PA (IMMP-) IMMPHERON INC.
XX
PI Kohler H, Muller S, Brown TL, Zhao Y, Morgan AC;
XX
DR WPI; 2004-653567/63.
XX
PT New compound for regulating normal or infected cell function comprising
PT an antibody conjugated to a membrane transporter peptide, useful in
PT preparing a composition for treating or preventing human diseases, e.g.
PT Alzheimer's disease.
XX
PS Example 1; SEQ ID NO 1; 50pp; English.
XX
CC The specification describes a fusion protein for regulating normal or
CC infected cell function, comprising an antibody conjugated to a peptide
CC having immunostimulatory, membrane transport, and homophilic activities.
CC The antibody is immunospecific for a signaling protein internal cell
CC consisting of caspases, kinases or phosphatases, an immature viral
CC protein, a cell-surface or intracellular tumour antigen, a nuclear or
CC nucleolar protein participating in regulation of DNA synthesis and gene
CC expression, or a cytoskeletal protein participating in cell proliferation
CC or cytostasis. The peptide portion of the fusion protein is preferably a
CC membrane transporter peptide that is endogenous to Kaposi fibroblast
CC factor, TAT peptides of HIV-1, antennapedia homeodomain-derived peptide,
CC herpes virus protein VP22, or transportan peptide. Fusion protein of the
CC invention are useful for preparing a composition for treating or
CC preventing human diseases, e.g., Alzheimer's disease, Huntington's
CC disease or Parkinson's disease. The present sequence represents a peptide
CC derived from the C3d region 1217-1232, which was affinity cross-linked
CC with 3H1 monoclonal antibody to produce fusion proteins of the invention.
CC 3H1 is a murine anti-idiotypic antibody which mimics the carcino-
CC embryonic antigen (CEA), and induces anti-CEA antibodies. The resulting
CC Cd3-3H1 fusion protein was used to enhance an anti-idiotype vaccine.
XX
SQ Sequence 16 AA;

```
  Query Match           100.0%;  Score 91;  DB 8;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 5.6e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          1 KNRWEDPGKQLYNVEA 16
```

RESULT 4
AAB71451
ID AAB71451 standard; peptide; 63 AA.
XX
AC AAB71451;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human C3 discontinous Factor H binding site SEQ ID 20.
XX
KW CVF1; cobra venom factor; proCVF1; cobra; antirheumatic; antiarthritic;
KW dermatological; immunosuppressive; vasotropic; vulnerary; septic shock;
KW antiinflammatory; antibacterial; decomplementation; rheumatoid arthritis;
KW ischaemia-reperfusion injury; thermal injury; lupus erythematosus;
KW respiratory distress syndrome; tissue rejection; complement; tumour;
KW xenotransplantation; gene therapy; burn; cancer.
XX
OS Homo sapiens.
XX
PN US2002103346-A1.
XX
PD 01-AUG-2002.
XX
PF 10-AUG-2001; 2001US-00925442.
XX
PR 14-JUN-1996; 96US-00662227.
PR 03-FEB-1998; 98US-00017947.
XX
PA (GEOU) UNIV GEORGETOWN.
XX
PI Vogel C, Bredehorst R, Fritzinger D, Kock M;
XX
DR WPI; 2002-690629/74.
XX
PT Novel recombinant pro-cobra venom factor polypeptide useful for
PT decomplementation of animal suffering from septic shock, ischemia-
PT reperfusion injury, arthritis, respiratory distress syndrome, or tissue
PT rejection.
XX
PS Example; Fig 7B; 82pp; English.
XX
CC This invention describes a novel recombinant pro-cobra venom factor
CC polypeptide which has antirheumatic, antiarthritic, dermatological,
CC immunosuppressive, vasotropic, vulnerary, antiinflammatory, antibacterial
CC and cytostatic activity. The polypeptide of the invention is useful for
CC decomplementation by administering proCVF to an animal such as reptile,
CC fish, bird or mammal such as guinea pigs, mice, rats, pigs, baboons,
CC chimps, dogs, cats, horses, cows or humans, suffering from septic shock,
CC ischaemia-reperfusion injury, thermal injury, arthritis, lupus,
CC respiratory distress syndrome, or a tissue rejection. ProCVF is useful as
CC research agent to deplete the complement activity in the plasma of
CC laboratory animals in vitro and in vivo, as therapeutic agent in humans
CC for treating cancer, for antibody targeting to tumour cells, for
CC depleting complement in patients undergoing xenotransplantation to
CC suppress the hyperacute rejection of the foreign organ, for temporary
CC depletion of complement in patient undergoing gene therapy using
CC retroviral vectors, and for treating diseases with circulating immune
CC complexes e.g. rheumatoid arthritis, lupus erythematosus, septic shock,
CC adult respiratory distress syndrome, ischaemic-reperfusion injury and
CC thermal injury from burns. This sequence represents a fragment of protein
CC described in the disclosure of the invention
XX
SQ Sequence 63 AA;

```
  Query Match           100.0%;  Score 91;  DB 5;  Length 63;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          9 KNRWEDPGKQLYNVEA 24
```

RESULT 5
AAU74858
ID AAU74858 standard; protein; 294 AA.
XX
AC AAU74858;
XX
DT 09-APR-2002 (first entry)
XX
DE Complement pathway protein C3d, R49A mutant.
XX
KW Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 49
FT /note= "Wild type Arg substituted by Ala"
XX
PN WO200192295-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-CA000785.
XX
PR 30-MAY-2000; 2000US-0207434P.
XX
PA (UTOR) UNIV TORONTO.
XX
PI Isenman DE, Clemenza L;
XX
DR WPI; 2002-114323/15.
XX
PT Ligand useful for modulating immune response such as in the preparation
PT of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS Disclosure; Page; 53pp; English.
XX
CC The invention describes a ligand of the complement receptor 2 (CD21 or
CC CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC molecule. The ligand is useful in the manufacture of a medicament such as
CC a vaccine for modulating the immune response of a host (preferably tumour
CC vaccine), and as antigens in immunogenic compositions, therapeutics
CC diagnostic reagents, in the generation of diagnostic agents and as cancer
CC therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC site specific mutated analogues of C3d demonstrate an enhanced binding
CC affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC inducing or enhancing an immune response to an antigen in a host and thus
CC protects the host against disease caused by the pathogen. This sequence
CC represents the complement pathway protein C3d R49A mutant, used to study
CC the interaction of C3d with complement receptor 2 (CD21/CD2), described
CC in the method of the invention. Note: This sequence does not appear in
CC the specification but has been created from a C3d wild type sequence
CC referenced on page 11 of the invention
XX
SQ Sequence 294 AA;

```
  Query Match           100.0%;  Score 91;  DB 5;  Length 294;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239
```

```
RESULT 6
AAU74866
ID   AAU74866 standard; protein; 294 AA.
XX
AC   AAU74866;
XX
DT   09-APR-2002  (first entry)
XX
DE   Complement pathway protein C3d, N98A mutant.
XX
KW   Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW   vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 98
FT                   /note= "Wild type Asn substituted by Ala"
XX
PN   WO200192295-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-CA000785.
XX
PR   30-MAY-2000; 2000US-0207434P.
XX
PA   (UTOR ) UNIV TORONTO.
XX
PI   Isenman DE,  Clemenza L;
XX
DR   WPI; 2002-114323/15.
XX
PT   Ligand useful for modulating immune response such as in the preparation
PT   of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS   Disclosure; Page; 53pp; English.
XX
CC   The invention describes a ligand of the complement receptor 2 (CD21 or
CC   CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC   molecule. The ligand is useful in the manufacture of a medicament such as
CC   a vaccine for modulating the immune response of a host (preferably tumour
CC   vaccine), and as antigens in immunogenic compositions, therapeutics
CC   diagnostic reagents, in the generation of diagnostic agents and as cancer
CC   therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC   cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC   site specific mutated analogues of C3d demonstrate an enhanced binding
CC   affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC   molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC   inducing or enhancing an immune response to an antigen in a host and thus
CC   protects the host against disease caused by the pathogen. This sequence
CC   represents the complement pathway protein C3d N98A mutant, used to study
CC   the interaction of C3d with complement receptor 2 (CD21/CD2), described
CC   in the method of the invention. Note: This sequence does not appear in
CC   the specification but has been created from a C3d wild type sequence
CC   referenced on page 11 of the invention
XX
SQ   Sequence 294 AA;

  Query Match             100.0%;  Score 91;  DB 5;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239
```

```
RESULT 7
AAU74869
ID   AAU74869 standard; protein; 294 AA.
XX
AC   AAU74869;
XX
DT   09-APR-2002  (first entry)
XX
DE   Complement pathway protein C3d, D163A mutant.
XX
KW   Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW   vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 163
FT                   /note= "Wild type Asp substituted by Ala"
XX
PN   WO200192295-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-CA000785.
XX
PR   30-MAY-2000; 2000US-0207434P.
XX
PA   (UTOR ) UNIV TORONTO.
XX
PI   Isenman DE,  Clemenza L;
XX
DR   WPI; 2002-114323/15.
XX
PT   Ligand useful for modulating immune response such as in the preparation
PT   of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS   Disclosure; Page; 53pp; English.
XX
CC   The invention describes a ligand of the complement receptor 2 (CD21 or
CC   CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC   molecule. The ligand is useful in the manufacture of a medicament such as
CC   a vaccine for modulating the immune response of a host (preferably tumour
CC   vaccine), and as antigens in immunogenic compositions, therapeutics
CC   diagnostic reagents, in the generation of diagnostic agents and as cancer
CC   therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC   cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC   site specific mutated analogues of C3d demonstrate an enhanced binding
CC   affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC   molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC   inducing or enhancing an immune response to an antigen in a host and thus
CC   protects the host against disease caused by the pathogen. This sequence
CC   represents the complement pathway protein C3d D163A mutant, used to study
CC   the interaction of C3d with complement receptor 2 (CD21/CD2), described
CC   in the method of the invention. Note: This sequence does not appear in
CC   the specification but has been created from a C3d wild type sequence
CC   referenced on page 11 of the invention
XX
SQ   Sequence 294 AA;

  Query Match             100.0%;  Score 91;  DB 5;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239
```

```
RESULT 8
AAU74855
```

```
ID   AAU74855 standard; protein; 294 AA.
XX
AC   AAU74855;
XX
DT   09-APR-2002  (first entry)
XX
DE   Complement pathway protein C3d, E37A mutant.
XX
KW   Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW   vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 37
FT                   /note= "Wild type Glu substituted by Ala"
XX
PN   WO200192295-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-CA000785.
XX
PR   30-MAY-2000; 2000US-0207434P.
XX
PA   (UTOR ) UNIV TORONTO.
XX
PI   Isenman DE,  Clemenza L;
XX
DR   WPI; 2002-114323/15.
XX
PT   Ligand useful for modulating immune response such as in the preparation
PT   of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS   Disclosure; Page; 53pp; English.
XX
CC   The invention describes a ligand of the complement receptor 2 (CD21 or
CC   CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC   molecule. The ligand is useful in the manufacture of a medicament such as
CC   a vaccine for modulating the immune response of a host (preferably tumour
CC   vaccine), and as antigens in immunogenic compositions, therapeutics
CC   diagnostic reagents, in the generation of diagnostic agents and as cancer
CC   therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC   cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC   site specific mutated analogues of C3d demonstrate an enhanced binding
CC   affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC   molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC   inducing or enhancing an immune response to an antigen in a host and thus
CC   protects the host against disease caused by the pathogen. This sequence
CC   represents the complement pathway protein C3d E37A mutant, used to study
CC   the interaction of C3d with complement receptor 2 (CD21/CD2), described
CC   in the method of the invention. Note: This sequence does not appear in
CC   the specification but has been created from a C3d wild type sequence
CC   referenced on page 11 of the invention
XX
SQ   Sequence 294 AA;

  Query Match             100.0%;  Score 91;  DB 5;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239


RESULT 9
AAU74862
ID   AAU74862 standard; protein; 294 AA.
XX
AC   AAU74862;
XX
DT   09-APR-2002  (first entry)
XX
DE   Complement pathway protein C3d, E37A/E39A mutant.
XX
KW   Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW   vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 37
FT                   /note= "Wild type Glu substituted by Ala"
FT   Misc-difference 39
FT                   /note= "Wild type Glu substituted by Ala"
XX
PN   WO200192295-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-CA000785.
XX
PR   30-MAY-2000; 2000US-0207434P.
XX
PA   (UTOR ) UNIV TORONTO.
XX
PI   Isenman DE,  Clemenza L;
XX
DR   WPI; 2002-114323/15.
XX
PT   Ligand useful for modulating immune response such as in the preparation
PT   of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS   Disclosure; Page; 53pp; English.
XX
CC   The invention describes a ligand of the complement receptor 2 (CD21 or
CC   CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC   molecule. The ligand is useful in the manufacture of a medicament such as
CC   a vaccine for modulating the immune response of a host (preferably tumour
CC   vaccine), and as antigens in immunogenic compositions, therapeutics
CC   diagnostic reagents, in the generation of diagnostic agents and as cancer
CC   therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC   cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC   site specific mutated analogues of C3d demonstrate an enhanced binding
CC   affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC   molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC   inducing or enhancing an immune response to an antigen in a host and thus
CC   protects the host against disease caused by the pathogen. This sequence
CC   represents the complement pathway protein C3d E37A/E39A mutant, used to
CC   study the interaction of C3d with complement receptor 2 (CD21/CD2),
CC   described in the method of the invention. Note: This sequence does not
CC   appear in the specification but has been created from a C3d wild type
CC   sequence referenced on page 11 of the invention
XX
SQ   Sequence 294 AA;

  Query Match             100.0%;  Score 91;  DB 5;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239


RESULT 10
AAU74859
ID   AAU74859 standard; protein; 294 AA.
XX
AC   AAU74859;
XX
DT   09-APR-2002  (first entry)
XX
```

DE Complement pathway protein C3d, R49M mutant.
XX
KW Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 49
FT /note= "Wild type Arg substituted by Met"
XX
PN WO200192295-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-CA000785.
XX
PR 30-MAY-2000; 2000US-0207434P.
XX
PA (UTOR) UNIV TORONTO.
XX
PI Isenman DE, Clemenza L;
XX
DR WPI; 2002-114323/15.
XX
PT Ligand useful for modulating immune response such as in the preparation
PT of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS Disclosure; Page; 53pp; English.
XX
CC The invention describes a ligand of the complement receptor 2 (CD21 or
CC CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC molecule. The ligand is useful in the manufacture of a medicament such as
CC a vaccine for modulating the immune response of a host (preferably tumour
CC vaccine), and as antigens in immunogenic compositions, therapeutics
CC diagnostic reagents, in the generation of diagnostic agents and as cancer
CC therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC site specific mutated analogues of C3d demonstrate an enhanced binding
CC affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC inducing or enhancing an immune response to an antigen in a host and thus
CC protects the host against disease caused by the pathogen. This sequence
CC represents the complement pathway protein C3d R49M mutant, used to study
CC the interaction of C3d with complement receptor 2 (CD21/CD2), described
CC in the method of the invention. Note: This sequence does not appear in
CC the specification but has been created from a C3d wild type sequence
CC referenced on page 11 of the invention
XX
SQ Sequence 294 AA;

Query Match 100.0%; Score 91; DB 5; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 224 KNRWEDPGKQLYNVEA 239

RESULT 11
AAU74872
ID AAU74872 standard; protein; 294 AA.
XX
AC AAU74872;
XX
DT 09-APR-2002 (first entry)
XX
DE Complement pathway protein C3d, E166A mutant.
XX
KW Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 166
FT /note= "Wild type Glu substituted by Ala"
XX
PN WO200192295-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-CA000785.
XX
PR 30-MAY-2000; 2000US-0207434P.
XX
PA (UTOR) UNIV TORONTO.
XX
PI Isenman DE, Clemenza L;
XX
DR WPI; 2002-114323/15.
XX
PT Ligand useful for modulating immune response such as in the preparation
PT of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS Disclosure; Page; 53pp; English.
XX
CC The invention describes a ligand of the complement receptor 2 (CD21 or
CC CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC molecule. The ligand is useful in the manufacture of a medicament such as
CC a vaccine for modulating the immune response of a host (preferably tumour
CC vaccine), and as antigens in immunogenic compositions, therapeutics
CC diagnostic reagents, in the generation of diagnostic agents and as cancer
CC therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC site specific mutated analogues of C3d demonstrate an enhanced binding
CC affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC inducing or enhancing an immune response to an antigen in a host and thus
CC protects the host against disease caused by the pathogen. This sequence
CC represents the complement pathway protein C3d E166A mutant, used to study
CC the interaction of C3d with complement receptor 2 (CD21/CD2), described
CC in the method of the invention. Note: This sequence does not appear in
CC the specification but has been created from a C3d wild type sequence
CC referenced on page 11 of the invention
XX
SQ Sequence 294 AA;

Query Match 100.0%; Score 91; DB 5; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 224 KNRWEDPGKQLYNVEA 239

RESULT 12
AAU74873
ID AAU74873 standard; protein; 294 AA.
XX
AC AAU74873;
XX
DT 09-APR-2002 (first entry)
XX
DE Complement pathway protein C3d, E167A mutant.
XX
KW Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX

```
FH   Key             Location/Qualifiers
FT   Misc-difference 167
FT                   /note= "Wild type Glu substituted by Ala"
XX
PN   WO200192295-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-CA000785.
XX
PR   30-MAY-2000; 2000US-0207434P.
XX
PA   (UTOR ) UNIV TORONTO.
XX
PI   Isenman DE,  Clemenza L;
XX
DR   WPI; 2002-114323/15.
XX
PT   Ligand useful for modulating immune response such as in the preparation
PT   of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS   Disclosure; Page; 53pp; English.
XX
CC   The invention describes a ligand of the complement receptor 2 (CD21 or
CC   CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC   molecule. The ligand is useful in the manufacture of a medicament such as
CC   a vaccine for modulating the immune response of a host (preferably tumour
CC   vaccine), and as antigens in immunogenic compositions, therapeutics
CC   diagnostic reagents, in the generation of diagnostic agents and as cancer
CC   therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC   cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC   site specific mutated analogues of C3d demonstrate an enhanced binding
CC   affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC   molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC   inducing or enhancing an immune response to an antigen in a host and thus
CC   protects the host against disease caused by the pathogen. This sequence
CC   represents the complement pathway protein C3d E167A mutant, used to study
CC   the interaction of C3d with complement receptor 2 (CD21/CD2), described
CC   in the method of the invention. Note: This sequence does not appear in
CC   the specification but has been created from a C3d wild type sequence
CC   referenced on page 11 of the invention
XX
SQ   Sequence 294 AA;

  Query Match             100.0%;  Score 91;  DB 5;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239


RESULT 13
AAU74863
ID   AAU74863 standard; protein; 294 AA.
XX
AC   AAU74863;
XX
DT   09-APR-2002  (first entry)
XX
DE   Complement pathway protein C3d, D36A/E37A/E39A mutant.
XX
KW   Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW   vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 36
FT                   /note= "Wild type Asp substituted by Ala"
FT   Misc-difference 37
FT                   /note= "Wild type Glu substituted by Ala"
FT   Misc-difference 39
FT                   /note= "Wild type Glu substituted by Ala"
XX
PN   WO200192295-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-CA000785.
XX
PR   30-MAY-2000; 2000US-0207434P.
XX
PA   (UTOR ) UNIV TORONTO.
XX
PI   Isenman DE,  Clemenza L;
XX
DR   WPI; 2002-114323/15.
XX
PT   Ligand useful for modulating immune response such as in the preparation
PT   of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS   Disclosure; Page; 53pp; English.
XX
CC   The invention describes a ligand of the complement receptor 2 (CD21 or
CC   CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC   molecule. The ligand is useful in the manufacture of a medicament such as
CC   a vaccine for modulating the immune response of a host (preferably tumour
CC   vaccine), and as antigens in immunogenic compositions, therapeutics
CC   diagnostic reagents, in the generation of diagnostic agents and as cancer
CC   therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC   cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC   site specific mutated analogues of C3d demonstrate an enhanced binding
CC   affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC   molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC   inducing or enhancing an immune response to an antigen in a host and thus
CC   protects the host against disease caused by the pathogen. This sequence
CC   represents the complement pathway protein C3d D36A/E37A/E39A mutant, used
CC   to study the interaction of C3d with complement receptor 2 (CD21/CD2),
CC   described in the method of the invention. Note: This sequence does not
CC   appear in the specification but has been created from a C3d wild type
CC   sequence referenced on page 11 of the invention
XX
SQ   Sequence 294 AA;

  Query Match             100.0%;  Score 91;  DB 5;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239


RESULT 14
AAU74856
ID   AAU74856 standard; protein; 294 AA.
XX
AC   AAU74856;
XX
DT   09-APR-2002  (first entry)
XX
DE   Complement pathway protein C3d, E39A mutant.
XX
KW   Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW   vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 39
FT                   /note= "Wild type Glu substituted by Ala"
XX
```

PN WO200192295-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-CA000785.
XX
PR 30-MAY-2000; 2000US-0207434P.
XX
PA (UTOR) UNIV TORONTO.
XX
PI Isenman DE, Clemenza L;
XX
DR WPI; 2002-114323/15.
XX
PT Ligand useful for modulating immune response such as in the preparation
PT of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS Disclosure; Page; 53pp; English.
XX
CC The invention describes a ligand of the complement receptor 2 (CD21 or
CC CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC molecule. The ligand is useful in the manufacture of a medicament such as
CC a vaccine for modulating the immune response of a host (preferably tumour
CC vaccine), and as antigens in immunogenic compositions, therapeutics
CC diagnostic reagents, in the generation of diagnostic agents and as cancer
CC therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC site specific mutated analogues of C3d demonstrate an enhanced binding
CC affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC inducing or enhancing an immune response to an antigen in a host and thus
CC protects the host against disease caused by the pathogen. This sequence
CC represents the complement pathway protein C3d E39A mutant, used to study
CC the interaction of C3d with complement receptor 2 (CD21/CD2), described
CC in the method of the invention. Note: This sequence does not appear in
CC the specification but has been created from a C3d wild type sequence
CC referenced on page 11 of the invention
XX
SQ Sequence 294 AA;

Query Match 100.0%; Score 91; DB 5; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 224 KNRWEDPGKQLYNVEA 239

RESULT 15
AAU74880
ID AAU74880 standard; protein; 294 AA.
XX
AC AAU74880;
XX
DT 09-APR-2002 (first entry)
XX
DE Complement pathway protein C3d, K291A mutant.
XX
KW Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW vaccine; CD21/CD19 complex; tumour; cancer; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 291
FT /note= "Wild type Lys substituted by Ala"
XX
PN WO200192295-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-CA000785.
XX
PR 30-MAY-2000; 2000US-0207434P.
XX
PA (UTOR) UNIV TORONTO.
XX
PI Isenman DE, Clemenza L;
XX
DR WPI; 2002-114323/15.
XX
PT Ligand useful for modulating immune response such as in the preparation
PT of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS Disclosure; Page; 53pp; English.
XX
CC The invention describes a ligand of the complement receptor 2 (CD21 or
CC CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC molecule. The ligand is useful in the manufacture of a medicament such as
CC a vaccine for modulating the immune response of a host (preferably tumour
CC vaccine), and as antigens in immunogenic compositions, therapeutics
CC diagnostic reagents, in the generation of diagnostic agents and as cancer
CC therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC site specific mutated analogues of C3d demonstrate an enhanced binding
CC affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC inducing or enhancing an immune response to an antigen in a host and thus
CC protects the host against disease caused by the pathogen. This sequence
CC represents the complement pathway protein C3d K291A mutant, used to study
CC the interaction of C3d with complement receptor 2 (CD21/CD2), described
CC in the method of the invention. Note: This sequence does not appear in
CC the specification but has been created from a C3d wild type sequence
CC referenced on page 11 of the invention
XX
SQ Sequence 294 AA;

Query Match 100.0%; Score 91; DB 5; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 224 KNRWEDPGKQLYNVEA 239

Search completed: August 24, 2005, 23:40:31
Job time : 166 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 24, 2005, 23:32:29 ; Search time 39 Seconds
(without alignments)
39.474 Million cell updates/sec

Title: US-09-865-281A-1
Perfect score: 91
Sequence: 1 KNRWEDPGKQLYNVEA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	100.0	1663	1	C3HU	complement C3 prec
2	80	87.9	1663	1	C3RT	complement C3 prec
3	79	86.8	726	2	A27602	complement C3 - ra
4	73	80.2	1663	1	C3MS	complement C3 prec
5	58	63.7	1666	1	C3GP	complement C3 prec
6	52	57.1	1358	2	B86241	hypothetical prote
7	44.5	48.9	490	2	D71401	probable selenium-
8	44	48.4	590	2	A44068	cell pattern forma
9	43	47.3	78	2	G82153	hypothetical prote
10	43	47.3	432	2	T05236	hypothetical prote
11	43	47.3	591	2	F95084	pyruvate oxidase [
12	43	47.3	591	2	B97952	pyruvate oxidase (
13	43	47.3	852	2	T12016	envelope glycoprot
14	43	47.3	1651	1	C3NJ	complement C3 prec
15	42.5	46.7	537	2	B90598	ABC transporter at
16	42	46.2	923	2	E83574	hypothetical prote
17	41	45.1	359	2	S45700	G-alpha-11 protein
18	41	45.1	538	2	E85438	step II splicing f
19	41	45.1	574	2	T16230	hypothetical prote
20	41	45.1	2166	2	G70163	hypothetical prote
21	41	45.1	5138	2	B96695	hypothetical prote
22	40.5	44.5	400	1	JC1428	ketol-acid reducto
23	40	44.0	166	2	A85077	hypothetical prote
24	40	44.0	274	2	S75320	hypothetical prote
25	40	44.0	290	2	G82360	diaminopimelate ep
26	40	44.0	311	2	C89894	methionyl-tRNA for
27	40	44.0	331	2	AI2972	two component sens
28	40	44.0	331	2	B98310	probable transmemb
29	40	44.0	359	1	RGHUGY	GTP-binding regula
30	40	44.0	432	2	B96515	hypothetical prote
31	40	44.0	434	2	C96515	hypothetical prote
32	40	44.0	630	2	JQ1670	polygalacturonase
33	40	44.0	647	2	G70733	probable htpG prot
34	40	44.0	1320	2	E59092	hypothetical prote
35	39.5	43.4	447	2	T07705	hypothetical prote
36	39	42.9	274	2	AE0468	diaminopimelate ep
37	39	42.9	302	2	H86271	protein F16A14.8 [
38	39	42.9	357	2	AF2796	lipoprotein [impor
39	39	42.9	363	2	AE3597	ABC transporter pe
40	39	42.9	364	2	D95364	hypothetical prote
41	39	42.9	371	2	F97575	hypothetical prote
42	39	42.9	400	2	AI0104	probable galactosi
43	39	42.9	427	2	JC4565	chitinase (EC 3.2.
44	39	42.9	428	2	T08576	phenylalanine-tRNA
45	39	42.9	606	2	G72282	hypothetical prote

ALIGNMENTS

RESULT 1
C3HU
complement C3 precursor [validated] - human
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit; C
C;Species: Homo sapiens (man)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004
C;Accession: A94065; A37999; A92187; A27603; A23435; A45830; B45830; A01257; A01258
R;de Bruijn, M.H.L.; Fey, G.H.
Proc. Natl. Acad. Sci. U.S.A. 82, 708-712, 1985
A;Title: Human complement component C3: cDNA coding sequence and derived primary structu
A;Reference number: A94065; MUID:85140166; PMID:2579379
A;Accession: A94065
A;Molecule type: mRNA
A;Residues: 1-1663 <DEB>
A;Cross-references: UNIPROT:P01024; GB:K02765; NID:g179664; PIDN:AAA85332.1; PID:g179665
R;Vik, D.P.; Amiguet, P.; Moffat, G.J.; Fey, M.; Amiguet-Barras, F.; Wetsel, R.A.; Tack,
Biochemistry 30, 1080-1085, 1991
A;Title: Structural features of the human C3 gene: intron/exon organization, transcriptic
A;Reference number: A37999; MUID:91113687; PMID:1703437
A;Contents: intron/exon structure of gene
A;Accession: A37999
A;Molecule type: DNA
A;Residues: 1-25 <VIK>
A;Cross-references: GB:M63423
A;Note: the authors translated the codon GGT for residue 6 as Leu, CCC for residue 7 as I
R;Hugli, T.E.
J. Biol. Chem. 250, 8293-8301, 1975
A;Title: Human anaphylatoxin (C3a) from the third component of complement.
A;Reference number: A92187; MUID:76069169; PMID:1238393
A;Accession: A92187
A;Molecule type: protein
A;Residues: 672-680,'N',682-699,'Q',701-748 <HUG>
R;Daoudaki, M.E.; Becherer, J.D.; Lambris, J.D.
J. Immunol. 140, 1577-1580, 1988
A;Title: A 34-amino acid peptide of the third component of complement mediates properdin
A;Reference number: A27603; MUID:88154452; PMID:3279119
A;Accession: A27603
A;Molecule type: protein
A;Residues: 1409-1563 <DAO>
R;Hellman, U.; Eggertsen, G.; Engstrom, A.; Sjoquist, J.
Biochem. J. 230, 353-361, 1985
A;Title: Amino acid sequence of the trypsin-generated C3d fragment from human complement
A;Reference number: A23435; MUID:86025442; PMID:3876831
A;Accession: A23435
A;Molecule type: protein
A;Residues: 1002-1012,'E',1014-1303 <HEL>
A;Note: sequence corresponding to residues 1072-1100 was not determined but was taken fr
R;Poznansky, M.C.; Clissold, P.M.; Lachmann, P.J.
J. Immunol. 143, 1254-1258, 1989
A;Title: The difference between human C3F and C3S results from a single amino acid change
3.
A;Reference number: A45830; MUID:89309808; PMID:2473125

A;Accession: A45830
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1212-1215,'N',1217-1223 <POZ>
A;Note: this is the C3S allele
A;Accession: B45830
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1212-1223 <PO2>
R;Dolmer, K.; Sottrup-Jensen, L.
FEBS Lett. 315, 85-90, 1993
A;Title: Disulfide bridges in human complement component C3b.
A;Reference number: S27041; MUID:93106233; PMID:8416818
A;Contents: annotation; disulfide bonds
C;Comment: The sequence shown is the C3 fast (C3F) allele, which is found mainly in Cauc
C;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
rnative-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Genetics:
A;Gene: GDB:C3
A;Cross-references: GDB:119044; OMIM:120700
A;Map position: 19p13.3-19p13.3
A;Note: contains 41 exons
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-667/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F;23-667,672-1663/Product: complement C3 #status predicted <CC3>
F;23-667,749-1663/Product: C3b #status predicted <C3B>
F;672-1663/Product: complement C3 alpha chain #status predicted <CC3A>
F;672-748/Product: C3a anaphylatoxin #status predicted <C3T>
F;749-1663/Product: C3b alpha' chain #status predicted <C3BA>
F;946-1303/Product: C3dk fragment #status predicted <CDK>
F;955-1303/Product: C3dg fragment #status predicted <CDG>
F;955-1001/Product: C3g fragment #status predicted <C3G>
F;1002-1303/Product: C3d fragment #status experimental <C3D>
F;1424-1457/Region: properdin binding
F;85,939/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;559-816,627-662,693-720,694-727,707-728,873-1513,1101-1158,1358-1489,1389-1458,1506-15
F;748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted
F;954-955/Cleavage site: Arg-Glu (complement factor I) #status predicted
F;1010-1013/Cross-link: thiolester (Cys-Gln) #status experimental
F;1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1320-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1617/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 91; DB 1; Length 1663;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 1217 KNRWEDPGKQLYNVEA 1232

RESULT 2
C3RT
complement C3 precursor - rat
N;Alternate names: 37K phospholipase A2 inhibitory protein
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Dec-1992 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: S15764; A54562; A01260; B35979; A35979; PN0567; PN0566; A32281; S08692
R;Misumi, Y.; Sohda, M.; Ikehara, Y.
Nucleic Acids Res. 18, 2178, 1990
A;Title: Nucleotide and deduced amino acid sequence of rat complement C3.
A;Reference number: S15764; MUID:90245672; PMID:2336397
A;Accession: S15764
A;Molecule type: mRNA
A;Residues: 1-1663 <MIS>
A;Cross-references: UNIPROT:P01026; EMBL:X52477; NID:g56953; PIDN:CAA36716.1; PID:g56954
R;Sundstrom, S.A.; Komm, B.S.; Ponce-de-Leon, H.; Yi, Z.; Teuscher, C.; Lyttle, C.R.
J. Biol. Chem. 264, 16941-16947, 1989
A;Title: Estrogen regulation of tissue-specific expression of complement C3.
A;Reference number: A54562; MUID:89380332; PMID:2674144
A;Accession: A54562
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 'P',1316-1595 <SUN>
A;Cross-references: GB:M29866; NID:g203200; PIDN:AAA40837.1; PID:g554423
R;Jacobs, J.W.; Rubin, J.S.; Hugli, T.E.; Bogardt, R.A.; Mariz, I.K.; Daniels, J.S.; Daug
Biochemistry 17, 5031-5038, 1978
A;Title: Purification, characterization, and amino acid sequence of rat anaphylatoxin (C3
A;Reference number: A01260; MUID:79062262; PMID:309768
A;Accession: A01260
A;Molecule type: protein
A;Residues: 671-703,'K',705-720,'KL',723-748 <JAC>
A;Note: three disulfide bonds are present
R;Suwa, Y.; Kudo, I.; Imaizumi, A.; Okada, M.; Kamimura, T.; Suzuki, Y.; Chang, H.W.; Ha
Proc. Natl. Acad. Sci. U.S.A. 87, 2395-2399, 1990
A;Title: Proteinaceous inhibitors of phospholipase A-2 purified from inflammatory sites i
A;Reference number: A35979; MUID:90207203; PMID:2320562
A;Accession: B35979
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'X',998-1005 <SUW>
A;Accession: A35979
A;Molecule type: protein
A;Residues: 'X',961-962,'P',964-969 <SU2>
R;Nakagawa, H.; Komorita, N.
Biochem. Biophys. Res. Commun. 194, 1181-1187, 1993
A;Title: Complement component C3-derived neutrophil chemotactic factors purified from exu
A;Reference number: PN0566; MUID:93356786; PMID:8352775
A;Accession: PN0567
A;Molecule type: protein
A;Residues: 568-592 <NAK>
A;Note: amino end of a C3-derived peptide designated exudate neutrophil chemotactic facto
A;Accession: PN0566
A;Molecule type: protein
A;Residues: 671-687 <NA2>
A;Note: amino end of peptide designated neutrophil chemotactic factor 1 and probably ider
R;Kuivanen, P.C.; Capulong, R.B.; Harkins, R.N.; DeSombre, E.R.
Biochem. Biophys. Res. Commun. 158, 898-905, 1989
A;Title: The estrogen-responsive 110K and 74K rat uterine secretory proteins are structur
A;Reference number: A32281; MUID:89149812; PMID:2645873
A;Accession: A32281
A;Molecule type: protein
A;Residues: 25-41 <KUI>
A;Experimental source: 17beta-estradiol-stimulated uterus of immature rat
A;Note: the authors treat this 74K uterine secretory protein, identical as far as sequenc
ent
C;Comment: Complement C3 contains two chains, formed by removal of four residues and link
alternative complement pathways, releases the C3a anaphylatoxin from the amino end of th
rnative-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign par
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by prot
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; chemotaxis; complement alternate pathway; complement pathway; gl
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-666/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F;25-666,671-1663/Product: complement C3 #status predicted <CC3>
F;25-666,749-1663/Product: complement C3b #status predicted <C3B>
F;671-1663/Product: complement C3 alpha chain #status predicted <CC3A>
F;671-748/Product: C3a anaphylatoxin #status experimental <C3T>
F;749-1663/Product: complement C3b alpha' chain #status predicted <C3BA>
F;946-1303/Product: C3dk fragment #status predicted <CDK>
F;1002-1303/Product: C3d fragment #status predicted <C3D>
F;1424-1457/Region: properdin binding
F;558-816,626-661,693-720,694-727,707-728,873-1513,1101-1158,1358-1489,1389-1458,1506-151
F;748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted

F;939,1617/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1010-1013/Cross-link: thiolester (Cys-Gln) #status predicted
F;1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1320-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted

Query Match 87.9%; Score 80; DB 1; Length 1663;
Best Local Similarity 81.2%; Pred. No. 0.00013;
Matches 13; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
:|||:||:|||||||
Db 1217 RNRWEEPGQQLYNVEA 1232

RESULT 3
A27602
complement C3 - rabbit (fragment)
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 15-Dec-1988 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: A27602
R;Kusano, M.; Choi, N.H.; Tomita, M.; Yamamoto, K.; Migita, S.; Sekiya, T.; Nishimura, S
Immunol. Invest. 15, 365-378, 1986
A;Title: Nucleotide sequence of cDNA and derived amino acid sequence of rabbit complemen
A;Reference number: A27602; MUID:87006907; PMID:3019881
A;Accession: A27602
A;Molecule type: mRNA
A;Residues: 1-726 <KUS>
A;Cross-references: UNIPROT:P12247; GB:M32434; NID:g164862; PIDN:AAA31190.1; PID:g164863
C;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
rnative-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;

Query Match 86.8%; Score 79; DB 2; Length 726;
Best Local Similarity 81.2%; Pred. No. 7.7e-05;
Matches 13; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
|||||:||::||||||
Db 280 KNRWEEPGQRLYNVEA 295

RESULT 4
C3MS
complement C3 precursor - mouse
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C;Accession: A92459; B92459; A92460; A93938; A21898; A54561; S16369; S16189; I49563; I49
R;Lundwall, A.; Wetsel, R.A.; Domdey, H.; Tack, B.F.; Fey, G.H.
J. Biol. Chem. 259, 13851-13856, 1984
A;Title: Structure of murine complement component C3: I. Nucleotide sequence of cloned c
A;Reference number: A92459; MUID:85054818; PMID:6548745
A;Accession: A92459
A;Molecule type: mRNA
A;Residues: 1-724 <LU1>
A;Cross-references: UNIPROT:P01027
A;Accession: B92459
A;Molecule type: DNA
A;Residues: 1-124 <LU2>
R;Wetsel, R.A.; Lundwall, A.; Davidson, F.; Gibson, T.; Tack, B.F.; Fey, G.H.
J. Biol. Chem. 259, 13857-13862, 1984
A;Title: Structure of murine complement component C3: II. Nucleotide sequence of cloned
A;Reference number: A92460; MUID:85054819; PMID:6094532
A;Accession: A92460
A;Molecule type: mRNA
A;Residues: 671-1663 <WET>
R;Domdey, H.; Wiebauer, K.; Kazmaier, M.; Muller, V.; Odink, K.; Fey, G.
Proc. Natl. Acad. Sci. U.S.A. 79, 7619-7623, 1982
A;Title: Characterization of the mRNA and cloned cDNA specifying the third component of r
A;Reference number: A93938; MUID:83117730; PMID:6961437
A;Contents: C3a
A;Accession: A93938
A;Molecule type: mRNA
A;Residues: 671-748 <DOM>
R;Sottrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Lonblad, P.B.; Jones, C.M.; Wierzbi
Proc. Natl. Acad. Sci. U.S.A. 82, 9-13, 1985
A;Title: Common evolutionary origin of alpha2-macroglobulin and complement components C3
A;Reference number: A21898; MUID:85113177; PMID:2578664
A;Accession: A21898
A;Molecule type: mRNA
A;Residues: 25-1663 <SOT>
R;Hamada, J.; Cavanaugh, P.G.; Miki, K.; Nicolson, G.L.
Cancer Res. 53, 4418-4423, 1993
A;Title: A paracrine migration-stimulating factor for metastatic tumor cells secreted by
A;Reference number: A54561; MUID:93373334; PMID:8364938
A;Accession: A54561
A;Molecule type: protein
A;Residues: 25-41;749-760 <HAM>
A;Experimental source: migration-stimulating factor purified from medium conditioned by n
R;Sato, T.; Hong, M.H.; Jin, C.H.; Ishimi, Y.; Udagawa, N.; Shinki, T.; Abe, E.; Suda, T
FEBS Lett. 285, 21-24, 1991
A;Title: The specific production of the third component of complement by osteoblastic cel
A;Reference number: S16189; MUID:91293304; PMID:2065778
A;Accession: S16369
A;Molecule type: protein
A;Residues: 25-31 <SAT>
A;Accession: S16189
A;Status: preliminary
A;Molecule type: protein
A;Residues: 671-677,'X',679-680 <SA2>
R;Fey, G.; Domdey, H.; Wiebauer, K.; Whitehead, A.S.; Odink, K.
Springer Semin. Immunopathol. 6, 119-147, 1983
A;Title: Structure and expression of the C3 gene.
A;Reference number: I49563; MUID:84045280; PMID:6356427
A;Accession: I49563
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 25-136,'Q',138-240 <FEY>
A;Cross-references: GB:M35659; NID:g192280; PIDN:AAA37339.1; PID:g192281
R;Fey, G.H.; Wiebauer, K.; Domdey, H.
Ann. N. Y. Acad. Sci. 421, 307-312, 1983
A;Title: Amino acid sequences of mouse complement C3 derived from nucleotide sequences o
A;Reference number: I49576; MUID:84201365; PMID:6609661
A;Accession: I49576
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 658-761 <RES>
A;Cross-references: GB:M33032; NID:g192391; PIDN:AAA37378.1; PID:g192392
C;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
rnative-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Genetics:
A;Introns: 27/2; 90/3
A;Note: the list of introns may be incomplete
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-666/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F;25-666,671-1663/Product: complement C3 #status predicted <CC3>
F;25-666,749-1663/Product: C3b #status predicted <C3B>
F;671-1663/Product: complement C3 alpha chain #status predicted <CC3A>
F;671-748/Product: C3a anaphylatoxin #status predicted <C3T>
F;749-1663/Product: C3b alpha' chain #status predicted <C3BA>
F;946-1303/Product: C3dk fragment #status predicted <CDK>
F;1002-1303/Product: C3d fragment #status predicted <C3D>

```
F;1424-1457/Region: properdin binding
F;559-816,626-661,693-720,694-727,707-728,873-1513,1101-1158,1358-1489,1389-1458,1506-15
F;748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted
F;939,1617/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1010-1013/Cross-link: thiolester (Cys-Gln) #status predicted
F;1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1320-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted

  Query Match             80.2%;  Score 73;  DB 1;  Length 1663;
  Best Local Similarity   75.0%;  Pred. No. 0.0018;
  Matches   12;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              :||||:| :|||||||
Db       1217 RNRWEEPDQQLYNVEA 1232


RESULT 5
C3GP
complement C3 precursor - guinea pig
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
C;Species: Cavia porcellus (guinea pig)
C;Date: 07-Feb-1992 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: A37156; S03375; A20342; D20342; C20342; A31222
R;Auerbach, H.S.; Burger, R.; Dodds, A.; Colten, H.R.
J. Clin. Invest. 86, 96-106, 1990
A;Title: Molecular basis of complement C3 deficiency in guinea pigs.
A;Reference number: A37156; MUID:90307998; PMID:1973176
A;Accession: A37156
A;Molecule type: mRNA
A;Residues: 1-1666 <AUE>
A;Cross-references: UNIPROT:P12387; GB:M34054; NID:g191262; PIDN:AAA37038.1; PID:g305335
R;Gerard, N.P.; Lively, M.O.; Gerard, C.
Protein Seq. Data Anal. 1, 473-478, 1988
A;Title: Amino acid sequence of guinea pig C3a anaphylatoxin.
A;Reference number: S03375; MUID:89113342; PMID:3064079
A;Accession: S03375
A;Molecule type: protein
A;Residues: 676-730,'N',732-752 <GER>
A;Experimental source: complement-activated guinea pig serum
A;Note: form isolated is inactive C3a anaphylatoxin and is missing the carboxyl-terminal
R;Thomas, M.L.; Tack, B.F.
Biochemistry 22, 942-947, 1983
A;Title: Identification and alignment of a thiol ester site in the third component of gu
A;Reference number: A90479; MUID:83178889; PMID:6838833
A;Accession: A20342
A;Molecule type: protein
A;Residues: 676-687 <TH1>
A;Accession: D20342
A;Molecule type: protein
A;Residues: 993-1012,1014-1017,'E',1019-1030,'Y' <TH2>
R;Goldberger, G.; Thomas, M.L.; Tack, B.F.; Williams, J.; Colten, H.R.; Abraham, G.N.
J. Biol. Chem. 256, 12617-12619, 1981
A;Title: NH2-terminal structure and cleavage of guinea pig pro-C3, the precursor of the
A;Reference number: A20342; MUID:82075767; PMID:6458605
A;Accession: C20342
A;Molecule type: protein
A;Residues: 23-38 <GOL>
C;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
 alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
rnative-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-671/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F;23-671,676-1666/Product: complement C3 #status predicted <CC3>
F;23-671,754-1666/Product: complement C3b #status predicted <C3B>
F;676-1666/Product: complement C3 alpha chain #status predicted <CC3A>
F;676-753/Product: C3a anaphylatoxin #status predicted <C3T>
F;754-1666/Product: complement C3b alpha' chain #status predicted <C3BA>
F;951-1308/Product: C3dk fragment #status predicted <CDK>
F;1007-1308/Product: C3d fragment #status predicted <C3D>
F;1429-1461/Region: properdin binding
F;557-821,630-666,698-725,699-732,712-733,878-1517,1106-1163,1363-1493,1394-1462,1510-151
F;753-754/Cleavage site: Arg-Ser (C3 convertase) #status predicted
F;944,1620/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1015-1018/Cross-link: thiolester (Cys-Gln) #status experimental
F;1308-1309/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1325-1326/Cleavage site: Arg-Ser (complement factor I) #status predicted

  Query Match             63.7%;  Score 58;  DB 1;  Length 1666;
  Best Local Similarity   62.5%;  Pred. No. 0.5;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              |||||:  ::||:|||
Db       1222 KNRWEEARQKLYSVEA 1237


RESULT 6
B86241
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86241
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: B86241
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1358 <STO>
A;Cross-references: UNIPROT:Q9SAC6; GB:AE005172; NID:g4874272; PIDN:AAD31337.1; GSPDB:GN
C;Genetics:
A;Map position: 1

  Query Match             57.1%;  Score 52;  DB 2;  Length 1358;
  Best Local Similarity   69.2%;  Pred. No. 3.9;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 RWEDPGKQLYNVE 15
              |||  |||:|| |
Db        223 RWERKGKQMYNPE 235


RESULT 7
D71401
probable selenium-binding protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
C;Accession: D71401
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirks
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Giele
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech,
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ansc
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thali
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: D71401
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
```

```
A;Residues: 1-490 <BEV>
A;Cross-references: UNIPROT:O23264; GB:Z97335; NID:g2244747; PID:g2244759
C;Genetics:
A;Map position: 4COP9-4G3845
C;Superfamily: Caenorhabditis elegans hypothetical protein Y37A1B.5

  Query Match             48.9%;  Score 44.5;  DB 2;  Length 490;
  Best Local Similarity   56.2%;  Pred. No. 22;
  Matches    9;  Conservative    1;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 KNRWEDPG-KQLYNVE 15
              ||||| ||   ||  :
Db        191 KNRWEKPGHSPLYGYD 206


RESULT 8
A44068
cell pattern formation-associated protein - Emericella nidulans
N;Alternate names: cell differentiation and spatial organization regulator stuA
C;Species: Emericella nidulans, Aspergillus nidulans
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A44068; S27413
R;Miller, K.Y.; Wu, J.; Miller, B.L.
Genes Dev. 6, 1770-1782, 1992
A;Title: StuA is required for cell pattern formation in Aspergillus.
A;Reference number: A44068; MUID:92387550; PMID:1516832
A;Accession: A44068
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-590 <MIL>
A;Cross-references: UNIPROT:P36011; EMBL:M83569; NID:g168095; PID:g168096
A;Note: sequence extracted from NCBI backbone (NCBIP:112625)
C;Genetics:
A;Introns: 92/1; 157/1; 201/2
C;Keywords: DNA binding; nucleus; transcription regulation

  Query Match             48.4%;  Score 44;  DB 2;  Length 590;
  Best Local Similarity   61.5%;  Pred. No. 33;
  Matches    8;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 WEDPGKQLYNVEA 16
              ||| |   | |||
Db        135 WEDEGSLCYQVEA 147


RESULT 9
G82153
hypothetical protein VC1802 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82153
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82153
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-78 <HEI>
A;Cross-references: UNIPROT:Q9KR44; GB:AE004257; GB:AE003852; NID:g9656326; PIDN:AAF9495
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1802
A;Map position: 1

  Query Match             47.3%;  Score 43;  DB 2;  Length 78;
  Best Local Similarity   87.5%;  Pred. No. 5.6;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DPGKQLYN 13
              ||||:|||
Db         11 DPGKELYN 18


RESULT 10
T05236
hypothetical protein F18A5.60 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05236
R;Bevan, M.; Weber, N.; Grueninger, D.; Schmidheini, T.; Bancroft, I.; Mewes, H.W.; Maye
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15405
A;Accession: T05236
A;Molecule type: DNA
A;Residues: 1-432 <BEV>
A;Cross-references: UNIPROT:Q9SVP5; EMBL:AL035528
A;Experimental source: cultivar Columbia; BAC clone F18A5
C;Genetics:
A;Map position: 4
A;Introns: 191/3; 223/3; 274/2; 389/1; 401/3
A;Note: F18A5.60

  Query Match             47.3%;  Score 43;  DB 2;  Length 432;
  Best Local Similarity   66.7%;  Pred. No. 34;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGK 9
              :||||||  :
Db        348 ENRWEDPSR 356


RESULT 11
F95084
pyruvate oxidase [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 16-Aug-2004
C;Accession: F95084
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heide
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, E
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95084
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-591 <KUR>
A;Cross-references: UNIPROT:Q54970; GB:AE005672; PIDN:AAK74871.1; PID:g14972205; GSPDB:GN
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0730
C;Superfamily: Acetolactate synthase, large subunit/pyruvate oxidase

  Query Match             47.3%;  Score 43;  DB 2;  Length 591;
  Best Local Similarity   46.7%;  Pred. No. 48;
  Matches    7;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVE 15
              ||::||  | |: |:
Db        475 KNKYEDTNKHLFGVD 489


RESULT 12
B97952
pyruvate oxidase (EC 1.2.3.3) [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: B97952
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.; Es
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Mc
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
```

```
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: B97952
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-591 <KUR>
A;Cross-references: UNIPROT:Q8DQJ4; GB:AE007317; PIDN:AAK99446.1; PID:g15458227; GSPDB:G
C;Genetics:
A;Gene: spxB
C;Keywords: oxidoreductase

  Query Match             47.3%;  Score 43;  DB 2;  Length 591;
  Best Local Similarity   46.7%;  Pred. No. 48;
  Matches    7;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVE 15
              ||::||  | |: |:
Db        475 KNKYEDTNKHLFGVD 489


RESULT 13
T12016
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T12016
R;McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A;Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S
A;Reference number: Z17379; MUID:98178716; PMID:9519894
A;Accession: T12016
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-852 <MCC>
A;Cross-references: UNIPROT:O41883; EMBL:U90934; NID:g2351783; PIDN:AAC59271.1; PID:g235
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein

  Query Match             47.3%;  Score 43;  DB 2;  Length 852;
  Best Local Similarity   54.5%;  Pred. No. 70;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              |||:: || :|
Db        421 NRWQEVGKAMY 431


RESULT 14
C3NJ
complement C3 precursor - monocled cobra
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
C;Species: Naja naja kaouthia, Naja naja siamensis (monocled cobra)
C;Date: 18-Jun-1993 #sequence_revision 07-Oct-1994 #text_change 17-Mar-2000
C;Accession: A46513
R;Fritzinger, D.C.; Petrella, E.C.; Connelly, M.B.; Bredehorst, R.; Vogel, C.W.
J. Immunol. 149, 3554-3562, 1992
A;Title: Primary structure of cobra complement component C3.
A;Reference number: A46513; MUID:93056528; PMID:1431125
A;Accession: A46513
A;Molecule type: mRNA
A;Residues: 1-1651 <FRI>
A;Cross-references: GB:L02365; NID:g213372; PIDN:AAA49385.1; PID:g213373
A;Note: authors' translation shows Arg-1408 after residue 1438 and, consequently, residu
A;Note: sequence extracted from NCBI backbone (NCBIP:118403) and corrected to correspond
C;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
 alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
rnative-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-655/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F;23-655,660-1651/Product: complement C3 #status predicted <CC3>
F;23-655,739-1651/Product: complement C3b #status predicted <C3B>
F;660-1651/Product: complement C3 alpha chain #status predicted <CC3A>
F;660-738/Product: C3a anaphylatoxin #status predicted <C3T>
F;739-1651/Product: complement C3b alpha' chain #status predicted <C3BA>
F;1412-1445/Region: properdin binding
F;546-807,615-650,683-710,684-717,697-718,863-1501,1091-1147,1346-1477,1377-1446,1494-149
F;738-739/Cleavage site: Arg-Ser (C3 convertase) #status predicted
F;999-1002/Cross-link: thiolester (Cys-Gln) #status predicted

  Query Match             47.3%;  Score 43;  DB 1;  Length 1651;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVE 15
              :||||:   : :|:|
Db       1204 RNRWEEYNARTHNIE 1218


RESULT 15
B90598
ABC transporter atp-binding protein [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 16-Aug-2004
C;Accession: B90598
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmo
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: B90598
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-537 <KUR>
A;Cross-references: UNIPROT:Q98PN2; GB:AL445566; PID:g14090105; PIDN:CAC13863.1; GSPDB:GN
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPU_6900
A;Genetic code: SGC3
C;Superfamily: ATP-binding cassette homology

  Query Match             46.7%;  Score 42.5;  DB 2;  Length 537;
  Best Local Similarity   56.2%;  Pred. No. 52;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    3;  Gaps    1;

Qy          3 RWE---DPGKQLYNVE 15
              :||    ||||: |||
Db        308 KWEINRVPGKQILNVE 323


Search completed: August 24, 2005, 23:44:13
Job time : 41 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 24, 2005, 23:20:49 ; Search time 169 Seconds
(without alignments)
48.481 Million cell updates/sec

Title: US-09-865-281A-1
Perfect score: 91
Sequence: 1 KNRWEDPGKQLYNVEA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	100.0	1663	1	CO3_HUMAN	P01024 homo sapien
2	80	87.9	1663	1	CO3_RAT	P01026 rattus norv
3	79	86.8	726	1	CO3_RABIT	P12247 oryctolagus
4	77	84.6	349	2	O46544	O46544 ovis aries
5	76	83.5	154	2	Q29289	Q29289 sus scrofa
6	76	83.5	310	2	Q9Z115	Q9z115 mesocricetu
7	73	80.2	303	2	Q693V9	Q693v9 bos taurus
8	73	80.2	1663	1	CO3_MOUSE	P01027 mus musculu
9	73	80.2	1663	2	Q80XP1	Q80xp1 mus musculu
10	71	78.0	1661	2	Q9GKP1	Q9gkp1 sus scrofa
11	58	63.7	1666	1	CO3_CAVPO	P12387 cavia porce
12	52	57.1	1358	2	Q9SAC6	Q9sac6 arabidopsis
13	52	57.1	1399	2	Q9FPP2	Q9fpp2 arabidopsis
14	52	57.1	1540	2	Q9SGX4	Q9sgx4 arabidopsis
15	50	54.9	92	2	Q9MXA7	Q9mxa7 barbus inte
16	48	52.7	922	2	Q8XIG1	Q8xig1 clostridium
17	47.5	52.2	373	2	Q9F2H7	Q9f2h7 staphylococ
18	47.5	52.2	401	2	Q8DYL9	Q8dyl9 streptococc
19	47.5	52.2	401	2	Q8E475	Q8e475 streptococc
20	46.5	51.1	462	2	Q839B3	Q839b3 enterococcu
21	46	50.5	92	2	Q9MXB8	Q9mxb8 barbus inte
22	46	50.5	151	2	P91717	P91717 dugesia tig
23	46	50.5	329	2	Q70XU5	Q70xu5 barbus inte
24	46	50.5	400	2	Q92270	Q92270 rhizopus ol
25	46	50.5	467	2	Q8VUW4	Q8vuw4 staphylococ
26	46	50.5	546	2	Q6AP19	Q6ap19 desulfotale
27	46	50.5	1475	2	Q8LPT9	Q8lpt9 citrus reti
28	45	49.5	172	2	Q7ZTW3	Q7ztw3 brachydanio
29	45	49.5	260	2	Q6NY31	Q6ny31 brachydanio
30	45	49.5	1684	2	Q9DDV9	Q9ddv9 oncorhynchu
31	44.5	48.9	490	1	SBP_ARATH	O23264 arabidopsis
32	44	48.4	163	2	O73133	O73133 human immun
33	44	48.4	205	2	Q8AL84	Q8al84 human immun
34	44	48.4	257	2	Q6VPT6	Q6vpt6 sarcoptes s
35	44	48.4	391	2	Q87DC6	Q87dc6 xylella fas
36	44	48.4	405	2	Q6FUX4	Q6fux4 candida gla
37	44	48.4	409	2	Q8S2S4	Q8s2s4 thellungiel
38	44	48.4	455	2	Q7ZBH7	Q7zbh7 simian-huma
39	44	48.4	473	2	O93081	O93081 human immun
40	44	48.4	478	2	O57038	O57038 human immun
41	44	48.4	499	2	Q7P1I4	Q7p1i4 chromobacte
42	44	48.4	622	1	STUA_EMENI	P36011 emericella
43	44	48.4	632	2	Q8NKF5	Q8nkf5 penicillium
44	44	48.4	698	2	Q7BSH7	Q7bsh7 rhizobium l
45	43	47.3	78	2	Q9KR44	Q9kr44 vibrio chol

ALIGNMENTS

RESULT 1
CO3_HUMAN
ID CO3_HUMAN STANDARD; PRT; 1663 AA.
AC P01024;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Complement C3 precursor [Contains: C3a anaphylatoxin].
GN Name=C3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85140166; PubMed=2579379;
RA de Bruijn M.H.L., Fey G.H.;
RT "Human complement component C3: cDNA coding sequence and derived
RT primary structure.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:708-712(1985).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS GLY-102; PRO-314; LYS-863; ASP-1224
RP AND THR-1367.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RT "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 672-748.
RX MEDLINE=76069169; PubMed=1238393;
RA Hugli T.E.;
RT "Human anaphylatoxin (C3a) from the third component of complement.
RT Primary structure.";
RL J. Biol. Chem. 250:8293-8301(1975).
RN [4]
RP SEQUENCE OF 955-966, AND SUBUNITS.
RC TISSUE=Serum;
RX MEDLINE=95293954; PubMed=7539791; DOI=10.1074/jbc.270.23.13645;
RA Oxvig C., Haaning J., Kristensen L., Wagner J.M., Rubin I.,
RA Stigbrand T., Gleich G.J., Sottrup-Jensen L.;
RT "Identification of angiotensinogen and complement C3dg as novel
RT proteins binding the proform of eosinophil major basic protein in
RT human pregnancy serum and plasma.";
RL J. Biol. Chem. 270:13645-13651(1995).
RN [5]
RP SEQUENCE OF 988-1036.
RX MEDLINE=82174534; PubMed=6175959;
RA Thomas M.L., Janatova J., Gray W.R., Tack B.F.;
RT "Third component of human complement: localization of the internal
RT thiolester bond.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1054-1058(1982).
RN [6]
RP SEQUENCE OF 1409-1563.

```
RX   MEDLINE=88154452; PubMed=3279119;
RA   Daoudaki M.E., Becherer J.D., Lambris J.D.;
RT   "A 34-amino acid peptide of the third component of complement mediates
RT   properdin binding.";
RL   J. Immunol. 140:1577-1580(1988).
RN   [7]
RP   STRUCTURE BY NMR OF C3A.
RX   MEDLINE=88276894; PubMed=3260670;
RA   Nettesheim D.G., Edalji R.P., Mollison K.W., Greer J.,
RA   Zuiderweg E.R.P.;
RT   "Secondary structure of complement component C3a anaphylatoxin in
RT   solution as determined by NMR spectroscopy: differences between
RT   crystal and solution conformations.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:5036-5040(1988).
RN   [8]
RP   MUTAGENESIS OF THIOESTER BOND REGION.
RX   MEDLINE=92250565; PubMed=1577777;
RA   Isaac L., Isenman D.E.;
RT   "Structural requirements for thioester bond formation in human
RT   complement component C3. Reassessment of the role of thioester bond
RT   integrity on the conformation of C3.";
RL   J. Biol. Chem. 267:10062-10069(1992).
RN   [9]
RP   DISULFIDE BONDS.
RX   MEDLINE=93106233; PubMed=8416818; DOI=10.1016/0014-5793(93)81139-Q;
RA   Dolmer K., Sottrup-Jensen L.;
RT   "Disulfide bridges in human complement component C3b.";
RL   FEBS Lett. 315:85-90(1993).
RN   [10]
RP   CARBOHYDRATE-LINKAGE SITE ASN-85.
RX   MEDLINE=22660472; PubMed=12754519; DOI=10.1038/nbt827;
RA   Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT   "Identification and quantification of N-linked glycoproteins using
RT   hydrazide chemistry, stable isotope labeling and mass spectrometry.";
RL   Nat. Biotechnol. 21:660-666(2003).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 996-1303.
RX   MEDLINE=98259089; PubMed=9596584; DOI=10.1126/science.280.5367.1277;
RA   Nagar B., Jones R.G., Diefenbach R.J., Isenman D.E., Rini J.M.;
RT   "X-ray crystal structure of C3d: a C3 fragment and ligand for
RT   complement receptor 2.";
RL   Science 280:1277-1281(1998).
RN   [12]
RP   VARIANT C3F/S.
RX   MEDLINE=89309808; PubMed=2473125;
RA   Poznansky M.C., Clissold P.M., Lachmann P.J.;
RT   "The difference between human C3F and C3S results from a single amino
RT   acid change from an asparagine to an aspartate residue at position
RT   1216 on the alpha-chain of the complement component, C3.";
RL   J. Immunol. 143:1254-1258(1989).
RN   [13]
RP   ERRATUM (RETRACTION).
RX   MEDLINE=90063087; PubMed=2584723;
RA   Poznansky M.C., Clissold P.M., Lachmann P.J.;
RL   J. Immunol. 143:3860-3862(1989).
RN   [14]
RP   VARIANTS GLY-102 AND PRO-314.
RX   MEDLINE=91011240; PubMed=1976733;
RA   Botto M., Yong Fong K., So A.K., Koch C., Walport M.J.;
RT   "Molecular basis of polymorphisms of human complement component C3.";
RL   J. Exp. Med. 172:1011-1017(1990).
RN   [15]
RP   VARIANT ASN-549.
RX   MEDLINE=95050640; PubMed=7961791;
RA   Singer L., Whitehead W.T., Akama H., Katz Y., Fishelson Z.,
RA   Wetsel R.A.;
RT   "Inherited human complement C3 deficiency. An amino acid substitution
RT   in the beta-chain (Asp549 to Asn) impairs C3 secretion.";
RL   J. Biol. Chem. 269:28494-28499(1994).
RN   [16]
RP   VARIANT GLN-1320.
RA   Watanabe Y., Matsui N., Yan K., Nishimukai H., Tokunaga K., Juji T.,
RA   Kobayashi N., Kohsaka T.;
RT   "A novel C3 allotype C3'F02'has an amino acid substitution that may
RT   inhibit iC3b synthesis and cause C3-hypocomplementemia.";
RL   Mol. Immunol. 30:62-62(1993).
CC   -!- FUNCTION: C3 plays a central role in the activation of the
CC       complement system. Its processing by C3 convertase is the central
CC       reaction in both classical and alternative complement pathways.
CC       After activation C3b can bind covalently, via its reactive
CC       thiolester, to cell surface carbohydrates or immune aggregates.
CC   -!- FUNCTION: Derived from proteolytic degradation of complement C3,
CC       C3a anaphylatoxin is a mediator of local inflammatory process. It
CC       induces the contraction of smooth muscle, increases vascular
CC       permeability and causes histamine release from mast cells and
CC       basophilic leukocytes.
CC   -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC       residues, forming two chains, beta and alpha, linked by a
CC       disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC       chain, releasing C3a anaphylatoxin and generating C3b (beta chain
CC       + alpha' chain). During pregnancy, C3dg exists as a complex
CC       (probably a 2:2:2 heterohexamer) with AGT and the proform of PRG2.
CC   -!- PTM: C3b is rapidly split in two positions by factor I and a
CC       cofactor to form iC3b (inactivated C3b) and C3f which is released.
CC       Then iC3b is slowly cleaved (possibly by factor I) to form C3c and
CC       C3dg. Other proteases produce other fragments such as C3d or C3g.
CC   -!- POLYMORPHISM: There are two alleles: C3S (C3 slow), the most
CC       common allele in all races and C3F (C3 fast), relatively frequent
CC       in Caucasoids, less common in Black Americans, extremely rare in
CC       Orientals.
CC   -!- DISEASE: Defects in C3 are the cause of C3 deficiency
CC       [MIM:120700]. It can result in susceptibility to pyogenic
CC       infection.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   -!- SIMILARITY: Contains 1 NTR domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; K02765; AAA85332.1; -.
DR   EMBL; AY513239; AAR89906.1; -.
DR   PIR; A94065; C3HU.
DR   PDB; 1C3D; X-ray; @=-.
DR   PDB; 1GHQ; X-ray; A=994-1300.
DR   GlycoSuiteDB; P01024; -.
DR   SWISS-2DPAGE; P01024; HUMAN.
DR   Siena-2DPAGE; P01024; -.
DR   Genew; HGNC:1318; C3.
DR   MIM; 120700; -.
DR   GO; GO:0005102; F:receptor binding; TAS.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR   GO; GO:0006955; P:immune response; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR008993; TIMP_like.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
```

```
DR   PROSITE; PS50189; NTR; 1.
KW   3D-structure; Complement alternate pathway; Complement pathway;

  Query Match             100.0%;  Score 91;  DB 1;  Length 1663;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db       1217 KNRWEDPGKQLYNVEA 1232


RESULT 2
CO3_RAT
ID   CO3_RAT        STANDARD;      PRT;  1663 AA.
AC   P01026;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Complement C3 precursor [Contains: C3a anaphylatoxin].
GN   Name=C3;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=90245672; PubMed=2336397;
RA   Misumi Y., Sohda M., Ikehara Y.;
RT   "Nucleotide and deduced amino acid sequence of rat complement C3.";
RL   Nucleic Acids Res. 18:2178-2178(1990).
RN   [2]
RP   SEQUENCE OF 671-748.
RX   MEDLINE=79062262; PubMed=309768;
RA   Jacobs J.W., Rubin J.S., Hugli T.E., Bogardt R.A., Mariz I.K.,
RA   Daniels J.S., Daughaday W.H., Bradshaw R.A.;
RT   "Purification, characterization, and amino acid sequence of rat
RT   anaphylatoxin (C3a).";
RL   Biochemistry 17:5031-5038(1978).
RN   [3]
RP   SEQUENCE OF 1316-1595 FROM N.A.
RX   MEDLINE=89380332; PubMed=2674144;
RA   Sundstrom S.A., Komm B.S., Ponce-De-Leon H., Yi Z., Teuscher C.,
RA   Lyttle C.R.;
RT   "Estrogen regulation of tissue-specific expression of complement C3.";
RL   J. Biol. Chem. 264:16941-16947(1989).
CC   -!- FUNCTION: C3 plays a central role in the activation of the
CC       complement system. Its processing by C3 convertase is the central
CC       reaction in both classical and alternative complement pathways.
CC       After activation C3b can bind covalently, via its reactive
CC       thiolester, to cell surface carbohydrates or immune aggregates.
CC   -!- FUNCTION: Derived from proteolytic degradation of complement C3,
CC       C3a anaphylatoxin is a mediator of local inflammatory process. It
CC       induces the contraction of smooth muscle, increases vascular
CC       permeability and causes histamine release from mast cells and
CC       basophilic leukocytes.
CC   -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC       residues, forming two chains, beta and alpha, linked by a
CC       disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC       chain, releasing C3a anaphylatoxin and generating C3b (beta chain
CC       + alpha' chain).
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   -!- SIMILARITY: Contains 1 NTR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52477; CAA36716.1; -.
DR   EMBL; M29866; AAA40837.1; ALT_SEQ.
DR   PIR; S15764; C3RT.
DR   PDB; 1QQF; X-ray; A=1010-1286.
DR   PDB; 1QSJ; X-ray; A/B/C/D=1010-1286.
DR   RGD; 2232; C3.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR008993; TIMP_like.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR   PROSITE; PS50189; NTR; 1.
KW   3D-structure; Complement alternate pathway; Complement pathway;
KW   Direct protein sequencing; Glycoprotein; Inflammatory response;
KW   Plasma; Signal; Thioester bond.
FT   SIGNAL       1     24
FT   CHAIN       25   1663       Complement C3.
FT   CHAIN       25    666       Complement C3 beta chain.
FT   CHAIN      671   1663       Complement C3 alpha chain.
FT   PEPTIDE    671    748       C3a anaphylatoxin.
FT   CHAIN      749   1663       Complement C3b alpha' chain.
FT   DOMAIN     693    728       Anaphylatoxin-like.
FT   DOMAIN    1518   1661       NTR.
FT   SITE       748    749       Cleavage (by C3 convertase).
FT   DISULFID   558    816       Interchain (By similarity).
FT   DISULFID   626    661       By similarity.
FT   DISULFID   693    720       By similarity.
FT   DISULFID   694    727       By similarity.
FT   DISULFID   707    728       By similarity.
FT   DISULFID   873   1513       By similarity.
FT   DISULFID  1101   1158       By similarity.
FT   DISULFID  1358   1489       By similarity.
FT   DISULFID  1389   1458       By similarity.
FT   DISULFID  1506   1511       By similarity.
FT   DISULFID  1518   1590       By similarity.
FT   DISULFID  1537   1661       By similarity.
FT   CROSSLNK  1010   1013       Isoglutamyl cysteine thioester (Cys-Gln).
FT   CARBOHYD   939    939       N-linked (GlcNAc. . .) (Probable).
FT   CARBOHYD  1617   1617       N-linked (GlcNAc. . .) (Probable).
FT   CONFLICT   721    722       LK -> KL (in Ref. 2).
FT   TURN      1011   1012
FT   HELIX     1013   1031
FT   TURN      1032   1032
FT   HELIX     1034   1037
FT   HELIX     1039   1041
FT   HELIX     1042   1057
FT   TURN      1058   1059
FT   STRAND    1060   1060
FT   TURN      1062   1063
FT   STRAND    1066   1066
FT   TURN      1070   1071
FT   HELIX     1076   1089
FT   TURN      1090   1092
FT   HELIX     1097   1111
FT   STRAND    1112   1112
FT   TURN      1114   1115
FT   STRAND    1118   1118
FT   HELIX     1127   1134
FT   TURN      1137   1138
FT   HELIX     1139   1158
```

```
FT   TURN       1159   1161
FT   TURN       1163   1164
FT   HELIX      1165   1180
FT   TURN       1181   1182
FT   HELIX      1186   1198
FT   TURN       1199   1200
FT   TURN       1204   1205
FT   HELIX      1206   1213
FT   STRAND     1215   1215
FT   TURN       1216   1218
FT   STRAND     1219   1219
FT   TURN       1223   1224
FT   HELIX      1226   1242
FT   TURN       1243   1244
FT   TURN       1246   1247
FT   HELIX      1249   1258
FT   TURN       1263   1264
FT   TURN       1266   1267
FT   HELIX      1269   1285
SQ   SEQUENCE   1663 AA;  186460 MW;  2F87CCB143CDD4BC CRC64;

  Query Match             87.9%;  Score 80;  DB 1;  Length 1663;
  Best Local Similarity   81.2%;  Pred. No. 0.0003;
  Matches   13;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              :||||:||:|||||||
Db       1217 RNRWEEPGQQLYNVEA 1232


RESULT 3
CO3_RABIT
ID   CO3_RABIT      STANDARD;      PRT;   726 AA.
AC   P12247;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Complement C3 alpha chain (Fragment).
GN   Name=C3;
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87006907; PubMed=3019881;
RA   Kusano M., Choi N.H., Tomita M., Yamamoto K., Migita S., Sekiya T.,
RA   Nishimura S.;
RT   "Nucleotide sequence of cDNA and derived amino acid sequence of rabbit
RT   complement component C3 alpha-chain.";
RL   Immunol. Invest. 15:365-378(1986).
CC   -!- FUNCTION: C3 plays a central role in the activation of the
CC       complement system. Its processing by C3 convertase is the central
CC       reaction in both classical and alternative complement pathways.
CC       After activation C3b can bind covalently, via its reactive
CC       thiolester, to cell surface carbohydrates or immune aggregates.
CC   -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC       residues, forming two chains, beta and alpha, linked by a
CC       disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC       chain, releasing C3a anaphylatoxin and generating C3b (beta chain
CC       + alpha' chain).
CC   -!- SIMILARITY: Contains 1 NTR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M32434; AAA31190.1; -.
DR   PIR; A27602; A27602.
DR   HSSP; P01024; 1C3D.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR008993; TIMP_like.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; PARTIAL.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; PARTIAL.
DR   PROSITE; PS50189; NTR; 1.
KW   Complement alternate pathway; Complement pathway; Glycoprotein;
KW   Inflammatory response; Plasma; Thioester bond.
FT   NON_TER       1      1
FT   CHAIN        <1    726       Complement C3 alpha chain.
FT   DOMAIN      581    724       NTR.
FT   CROSSLNK     73     76       Isoglutamyl cysteine thioester (Cys-Gln).
FT   CARBOHYD      2      2       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    233    233       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    680    680       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   726 AA;  81844 MW;  F4B4C35D461300E9 CRC64;

  Query Match             86.8%;  Score 79;  DB 1;  Length 726;
  Best Local Similarity   81.2%;  Pred. No. 0.00018;
  Matches   13;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              |||||:||::||||||
Db        280 KNRWEEPGQRLYNVEA 295


RESULT 4
O46544
ID   O46544      PRELIMINARY;      PRT;   349 AA.
AC   O46544;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Complement component C3 (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=white alpine; TISSUE=Liver;
RX   MEDLINE=98309471; PubMed=9647256;
RA   Hein W.R., Dudler L., Marston W.L., Landsverk T., Young A.J.,
RA   Avila D.;
RT   "Ubiquitination and dimerization of complement receptor type 2 on
RT   sheep B cells.";
RL   J. Immunol. 161:458-466(1998).
DR   EMBL; AF038130; AAB92374.2; -.
DR   HSSP; P01026; 1QQF.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
FT   NON_TER       1      1
FT   NON_TER     349    349
SQ   SEQUENCE   349 AA;  39679 MW;  70C2023E42ED5EE3 CRC64;

  Query Match             84.6%;  Score 77;  DB 2;  Length 349;
  Best Local Similarity   81.2%;  Pred. No. 0.00018;
  Matches   13;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              |||||:| |:||||||
Db        241 KNRWEEPNKKLYNVEA 256
```

```
RESULT 5
Q29289
ID   Q29289      PRELIMINARY;      PRT;   154 AA.
AC   Q29289;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Complement C3 (Fragment).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Small intestine;
RX   MEDLINE=96327607; PubMed=8672129;
RA   Winteroe A.K., Fredholm M., Davies W.;
RT   "Evaluation and characterization of a porcine small intestine cDNA
RT   library.";
RL   Mamm. Genome 7:509-517(1996).
DR   EMBL; F14640; CAA23173.1; -.
DR   HSSP; P01026; 1QQF.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   InterPro; IPR008930; Terp_cyc_toroid.
FT   NON_TER       1      1
FT   NON_TER     154    154
SQ   SEQUENCE   154 AA;  17440 MW;  6DC7661C1253ED45 CRC64;

  Query Match             83.5%;  Score 76;  DB 2;  Length 154;
  Best Local Similarity   75.0%;  Pred. No. 0.00011;
  Matches   12;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              :||||:||::||||||
Db         10 RNRWEEPGQKLYNVEA 25


RESULT 6
Q9Z115
ID   Q9Z115      PRELIMINARY;      PRT;   310 AA.
AC   Q9Z115;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Complement C3 (Fragment).
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yamamoto K., Inoue N., Sakiyama H.;
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB024425; BAA75923.1; -.
DR   HSSP; P01026; 1QQF.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
FT   NON_TER       1      1
FT   NON_TER     310    310
SQ   SEQUENCE   310 AA;  34779 MW;  11ED3BEEF82D327D CRC64;

  Query Match             83.5%;  Score 76;  DB 2;  Length 310;
  Best Local Similarity   75.0%;  Pred. No. 0.00023;
  Matches   12;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              :||||:||::||||||
Db        224 RNRWEEPGQKLYNVEA 239


RESULT 7
Q693V9
ID   Q693V9      PRELIMINARY;      PRT;   303 AA.
AC   Q693V9;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Complement component C3d (Fragment).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hodgins D., Firth M., Pei Y., Yoo D., Shewen P.;
RT   "Cloning, Sequencing and Analysis of the C3d Fragment of Bovine
RT   Complement Component 3.";
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY630404; AAT76518.1; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
FT   NON_TER       1      1
FT   NON_TER     303    303
SQ   SEQUENCE   303 AA;  34443 MW;  2F3A15020CEA3797 CRC64;

  Query Match             80.2%;  Score 73;  DB 2;  Length 303;
  Best Local Similarity   75.0%;  Pred. No. 0.00073;
  Matches   12;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              |||||:| ::||||||
Db        216 KNRWEEPNQKLYNVEA 231


RESULT 8
CO3_MOUSE
ID   CO3_MOUSE      STANDARD;      PRT;  1663 AA.
AC   P01027;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Complement C3 precursor (HSE-MSF) [Contains: C3a anaphylatoxin].
GN   Name=C3;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RX   MEDLINE=85038854; PubMed=6208565;
RA   Fey G.H., Lundwall A., Wetsel R.A., Tack B.F., de Bruijn M.H.L.,
RA   Domdey H.;
RT   "Nucleotide sequence of complementary DNA and derived amino acid
RT   sequence of murine complement protein C3.";
RL   Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:333-344(1984).
RN   [2]
RP   SEQUENCE OF 671-1663 FROM N.A. (ISOFORM LONG).
RX   MEDLINE=85054819; PubMed=6094532;
RA   Wetsel R.A., Lundwall A., Davidson F., Gibson T., Tack B.F., Fey G.H.;
RT   "Structure of murine complement component C3. II. Nucleotide sequence
RT   of cloned complementary DNA coding for the alpha chain.";
RL   J. Biol. Chem. 259:13857-13862(1984).
RN   [3]
RP   SEQUENCE OF 671-748 FROM N.A.
RX   MEDLINE=83117730; PubMed=6961437;
RA   Domdey H., Wiebauer K., Kazmaier M., Mueller V., Odink K., Fey G.H.;
RT   "Characterization of the mRNA and cloned cDNA specifying the third
RT   component of mouse complement.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:7619-7623(1982).
```

```
RN   [4]
RP   SEQUENCE OF 658-761 FROM N.A.
RX   MEDLINE=84201365; PubMed=6609661;
RA   Fey G.H., Wiebauer K., Domdey H.;
RT   "Amino acid sequences of mouse complement C3 derived from nucleotide
RT   sequences of cloned cDNA.";
RL   Ann. N. Y. Acad. Sci. 421:307-312(1983).
RN   [5]
RP   SEQUENCE OF 1-34 FROM N.A.
RX   MEDLINE=83117622; PubMed=6985486;
RA   Wiebauer K., Domdey H., Diggelmann H., Fey G.;
RT   "Isolation and analysis of genomic DNA clones encoding the third
RT   component of mouse complement.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:7077-7081(1982).
RN   [6]
RP   SEQUENCE OF 25-41 AND 749-760.
RX   MEDLINE=93373334; PubMed=8364938;
RA   Hamada J.-I., Cavanaugh P.G., Miki K., Nicolson G.L.;
RT   "A paracrine migration-stimulating factor for metastatic tumor cells
RT   secreted by mouse hepatic sinusoidal endothelial cells: identification
RT   as complement component C3b.";
RL   Cancer Res. 53:4418-4423(1993).
RN   [7]
RP   ALTERNATIVE INITIATION.
RX   MEDLINE=95053742; PubMed=7964485;
RA   Cahen-Kramer Y., Martensson I.L., Melchers F.;
RT   "The structure of an alternate form of complement C3 that displays
RT   costimulatory growth factor activity for B lymphocytes.";
RL   J. Exp. Med. 180:2079-2088(1994).
CC   -!- FUNCTION: C3 plays a central role in the activation of the
CC       complement system. Its processing by C3 convertase is the central
CC       reaction in both classical and alternative complement pathways.
CC       After activation C3b can bind covalently, via its reactive
CC       thiolester, to cell surface carbohydrates or immune aggregates.
CC   -!- FUNCTION: Derived from proteolytic degradation of complement C3,
CC       C3a anaphylatoxin is a mediator of local inflammatory process. It
CC       induces the contraction of smooth muscle, increases vascular
CC       permeability and causes histamine release from mast cells and
CC       basophilic leukocytes. The short isoform has B-cell stimulatory
CC       activity.
CC   -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC       residues, forming two chains, beta and alpha, linked by a
CC       disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC       chain, releasing C3a anaphylatoxin and generating C3b (beta chain
CC       + alpha' chain).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative initiation;
CC         Comment=2 isoforms, Long (shown here) and Short, are produced by
CC         alternative initiation;
CC   -!- PTM: C3b is rapidly split in two positions by factor I and a
CC       cofactor to form iC3b (inactivated C3b) and C3f which is released.
CC       Then iC3b is slowly cleaved (possibly by factor I) to form C3c and
CC       C3dg. Other proteases produce other fragments such as C3d or C3g.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   -!- SIMILARITY: Contains 1 NTR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02782; AAC42013.1; -.
DR   EMBL; J00369; AAA37336.1; -.
DR   EMBL; J00367; AAA37336.1; JOINED.
DR   EMBL; M33032; AAA37378.1; -.
DR   EMBL; Z37998; CAA86099.2; -.
DR   PIR; A92459; C3MS.
DR   HSSP; P01026; 1QQF.
DR   MGD; MGI:88227; C3.
DR   GO; GO:0006954; P:inflammatory response; IMP.
DR   GO; GO:0050766; P:positive regulation of phagocytosis; IMP.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR008993; TIMP_like.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR   PROSITE; PS50189; NTR; 1.
KW   Alternative initiation; Complement alternate pathway;
KW   Complement pathway; Direct protein sequencing; Glycoprotein;
KW   Inflammatory response; Plasma; Signal; Thioester bond.
FT   SIGNAL       1     24
FT   CHAIN       25   1663       Complement C3, isoform Long.
FT   CHAIN     1129   1663       Complement C3, isoform Short.
FT   INIT_MET  1129   1129       For isoform Short.
FT   CHAIN       25    666       Complement C3 beta chain.
FT   CHAIN      671   1663       Complement C3 alpha chain.
FT   PEPTIDE    671    748       C3a anaphylatoxin.
FT   CHAIN      749   1663       Complement C3b alpha' chain.
FT   CHAIN      749    954       Complement C3c fragment.
FT   CHAIN      955   1303       Complement C3dg fragment.
FT   CHAIN      955   1001       Complement C3g fragment.
FT   CHAIN     1002   1303       Complement C3d fragment.
FT   PEPTIDE   1304   1320       C3f fragment.
FT   DOMAIN     693    728       Anaphylatoxin-like.
FT   DOMAIN    1518   1661       NTR.
FT   SITE       748    749       Cleavage (by C3 convertase).
FT   SITE      1303   1304       Cleavage (by factor I).
FT   SITE      1320   1321       Cleavage (by factor I).
FT   DISULFID   559    816       Interchain (By similarity).
FT   DISULFID   626    661       By similarity.
FT   DISULFID   693    720       By similarity.
FT   DISULFID   694    727       By similarity.
FT   DISULFID   707    728       By similarity.
FT   DISULFID   873   1513       By similarity.
FT   DISULFID  1101   1158       By similarity.
FT   DISULFID  1358   1489       By similarity.
FT   DISULFID  1389   1458       By similarity.
FT   DISULFID  1506   1511       By similarity.
FT   DISULFID  1518   1590       By similarity.
FT   DISULFID  1537   1661       By similarity.
FT   DISULFID  1637   1646       By similarity.
FT   CARBOHYD   939    939       N-linked (GlcNAc. . .).
FT   CARBOHYD  1617   1617       N-linked (GlcNAc. . .).
FT   CROSSLNK  1010   1013       Isoglutamyl cysteine thioester (Cys-Gln)
FT                               (By similarity).
SQ   SEQUENCE   1663 AA;  186482 MW;  DE5546CC769BEA19 CRC64;

  Query Match             80.2%;  Score 73;  DB 1;  Length 1663;
  Best Local Similarity   75.0%;  Pred. No. 0.0046;
  Matches   12;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              :||||:| :|||||||
Db       1217 RNRWEEPDQQLYNVEA 1232


RESULT 9
Q80XP1
ID   Q80XP1      PRELIMINARY;      PRT;  1663 AA.
AC   Q80XP1;
```

```
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Complement component 3.
GN   Name=C3;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC043338; AAH43338.1; -.
DR   HSSP; P01026; 1QQF.
DR   MGD; MGI:88227; C3.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0050766; P:positive regulation of phagocytosis; IMP.
DR   GO; GO:0001798; P:positive regulation of type IIa hypersensit. . .; IMP.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR001599; MacrogloblnA2.
DR   InterPro; IPR001134; Netrin_C.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR008993; TIMP_like.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR   PROSITE; PS50189; NTR; 1.
SQ   SEQUENCE   1663 AA;  186483 MW;  7E5546CC7C314779 CRC64;

  Query Match             80.2%;  Score 73;  DB 2;  Length 1663;
  Best Local Similarity   75.0%;  Pred. No. 0.0046;
  Matches   12;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              :||||:| :|||||||
Db       1217 RNRWEEPDQQLYNVEA 1232


RESULT 10
Q9GKP1
ID   Q9GKP1      PRELIMINARY;      PRT;  1661 AA.
AC   Q9GKP1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Complement component C3 (Complement C3).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=21313047; PubMed=11419349;
RA   Wimmers K., Mekchay S., Ponsuksili S., Hardge T., Yerle M.,
RA   Schellander K.;
RT   "Polymorphic sites in exon 15 and 30 of the porcine C3 gene.";
RL   Anim. Genet. 32:46-47(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Wimmers K., Ponsuksili S., Schmoll F., Schellander K.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=22444329; PubMed=12557058;
RA   Wimmers K., Mekchay S., Schellander K., Ponsuksili S.;
RT   "Molecular characterization of the pig C3 gene and its association
RT   with complement activity.";
RL   Immunogenetics 54:714-724(2003).
DR   EMBL; AF154933; AAG40565.1; -.
DR   EMBL; AJ494748; CAD38823.2; -.
DR   HSSP; P01026; 1QQF.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0006956; P:complement activation; IEA.
DR   GO; GO:0006954; P:inflammatory response; IEA.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR001599; MacrogloblnA2.
DR   InterPro; IPR001134; Netrin_C.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR008993; TIMP_like.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR   PROSITE; PS50189; NTR; 1.
SQ   SEQUENCE   1661 AA;  186805 MW;  4899D0914BE3310C CRC64;

  Query Match             78.0%;  Score 71;  DB 2;  Length 1661;
  Best Local Similarity   68.8%;  Pred. No. 0.0099;
  Matches   11;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              :||||:||::|:||||
Db       1215 RNRWEEPGQKLHNVEA 1230
```

```
RESULT 11
CO3_CAVPO
ID   CO3_CAVPO      STANDARD;      PRT;  1666 AA.
AC   P12387;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Complement C3 precursor [Contains: C3a anaphylatoxin].
GN   Name=C3;
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90307998; PubMed=1973176;
RA   Auerbach H.S., Burger R., Dodds A., Colten H.R.;
RT   "Molecular basis of complement C3 deficiency in guinea pigs.";
RL   J. Clin. Invest. 86:96-106(1990).
RN   [2]
RP   SEQUENCE OF 676-753.
RX   MEDLINE=89113342; PubMed=3064079;
RA   Gerard N.P., Lively M.O., Gerard C.;
RT   "Amino acid sequence of guinea pig C3a anaphylatoxin.";
RL   Protein Seq. Data Anal. 1:473-478(1988).
RN   [3]
RP   SEQUENCE OF 993-1032.
RX   MEDLINE=83178889; PubMed=6838833;
RA   Thomas M.L., Tack B.F.;
RT   "Identification and alignment of a thiol ester site in the third
RT   component of guinea pig complement.";
RL   Biochemistry 22:942-947(1983).
CC   -!- FUNCTION: C3 plays a central role in the activation of the
CC       complement system. Its processing by C3 convertase is the central
CC       reaction in both classical and alternative complement pathways.
CC       After activation C3b can bind covalently, via its reactive
CC       thiolester, to cell surface carbohydrates or immune aggregates.
CC   -!- FUNCTION: Derived from proteolytic degradation of complement C3,
CC       C3a anaphylatoxin is a mediator of local inflammatory process. It
CC       induces the contraction of smooth muscle, increases vascular
CC       permeability and causes histamine release from mast cells and
CC       basophilic leukocytes.
CC   -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC       residues, forming two chains, beta and alpha, linked by a
CC       disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC       chain, releasing C3a anaphylatoxin and generating C3b (beta chain
CC       + alpha' chain).
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   -!- SIMILARITY: Contains 1 NTR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M34054; AAA37038.1; -.
DR   PIR; A37156; C3GP.
DR   HSSP; P01026; 1QQF.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR001599; MacrogloblnA2.
DR   InterPro; IPR001134; Netrin_C.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR008993; TIMP_like.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR   PROSITE; PS50189; NTR; 1.
KW   Complement alternate pathway; Complement pathway;
KW   Direct protein sequencing; Glycoprotein; Inflammatory response;
KW   Plasma; Signal; Thioester bond.
FT   SIGNAL        1     29
FT   CHAIN        30   1666       Complement C3.
FT   CHAIN        30    671       Complement C3 beta chain.
FT   CHAIN       676   1666       Complement C3 alpha chain.
FT   PEPTIDE     676    753       C3a anaphylatoxin.
FT   CHAIN       754   1666       Complement C3b alpha' chain.
FT   DOMAIN      698    733       Anaphylatoxin-like.
FT   DOMAIN     1522   1664       NTR.
FT   SITE        753    754       Cleavage (by C3 convertase).
FT   DISULFID    557    821       Interchain (By similarity).
FT   DISULFID    630    666       By similarity.
FT   DISULFID    698    725       By similarity.
FT   DISULFID    699    732       By similarity.
FT   DISULFID    712    733       By similarity.
FT   DISULFID    878   1517       By similarity.
FT   DISULFID   1106   1163       By similarity.
FT   DISULFID   1363   1493       By similarity.
FT   DISULFID   1394   1462       By similarity.
FT   DISULFID   1510   1515       By similarity.
FT   DISULFID   1522   1593       By similarity.
FT   DISULFID   1540   1664       By similarity.
FT   DISULFID   1640   1649       By similarity.
FT   CROSSLNK   1015   1018       Isoglutamyl cysteine thioester (Cys-Gln).
FT   CARBOHYD    944    944       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1620   1620       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT    731    731       D -> N (in Ref. 2).
FT   CONFLICT   1013   1013       Missing (in Ref. 3).
FT   CONFLICT   1018   1018       Q -> E (in Ref. 2).
FT   CONFLICT   1031   1031       Missing (in Ref. 3).
SQ   SEQUENCE   1666 AA;  186487 MW;  1C1F1219944AFD49 CRC64;

  Query Match             63.7%;  Score 58;  DB 1;  Length 1666;
  Best Local Similarity   62.5%;  Pred. No. 1.6;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              |||||:  ::||:|||
Db       1222 KNRWEEARQKLYSVEA 1237


RESULT 12
Q9SAC6
ID   Q9SAC6      PRELIMINARY;      PRT;  1358 AA.
AC   Q9SAC6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   T16B5.10 protein.
GN   Name=T16B5.10;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA   Lee J., Liu A., Li J., Kremenetskaia I., Luros J., Gonzalez A.,
RA   Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P.,
RA   Hansen N., Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
RA   Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Theologis;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC007354; AAD31337.1; -.
DR   PIR; B86241; B86241.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0016310; P:phosphorylation; IEA.
DR   InterPro; IPR002192; PPDK_N_term.
DR   Pfam; PF01326; PPDK_N; 1.
SQ   SEQUENCE   1358 AA;  151970 MW;  AA00A8D35163C3A5 CRC64;

  Query Match             57.1%;  Score 52;  DB 2;  Length 1358;
  Best Local Similarity   69.2%;  Pred. No. 13;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 RWEDPGKQLYNVE 15
              ||| |||:|| |
Db        223 RWERKGKQMYNPE 235


RESULT 13
Q9FPP2
ID   Q9FPP2      PRELIMINARY;      PRT;  1399 AA.
AC   Q9FPP2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   SEX1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21380420; PubMed=11487701;
RA   Yu T.-S., Kofler H., Hausler R.E., Hille D., Flugge U.-I.,
RA   Zeeman S.C., Smith A.M., Kossmann J., Lloyd J., Ritte G., Steup M.,
RA   Lue W.-L., Chen J., Weber A.;
RT   "The Arabidopsis sex1 mutant is defective in the R1 protein, a general
RT   regulator of starch degradation in plants, and not in the chloroplast
RT   hexose transporter.";
RL   Plant Cell 13:1907-1918(2001).
DR   EMBL; AF312027; AAG47821.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0016310; P:phosphorylation; IEA.
DR   InterPro; IPR002192; PPDK_N_term.
DR   Pfam; PF01326; PPDK_N; 1.
SQ   SEQUENCE   1399 AA;  156580 MW;  1FE9285376B479EB CRC64;

  Query Match             57.1%;  Score 52;  DB 2;  Length 1399;
  Best Local Similarity   69.2%;  Pred. No. 14;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 RWEDPGKQLYNVE 15
              ||| |||:|| |
Db        223 RWERKGKQMYNPE 235


RESULT 14
Q9SGX4
ID   Q9SGX4      PRELIMINARY;      PRT;  1540 AA.
AC   Q9SGX4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   F20B24.19.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA   Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA   Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA   Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA   Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA   Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA   Ecker J.R.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA   Shinn P., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA   Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA   Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
RA   Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA   Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA   Ecker J.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA   Shinn P., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA   Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA   Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
RA   Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA   Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA   Ecker J.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC009398; AAF17665.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004857; F:enzyme inhibitor activity; IEA.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0030599; F:pectinesterase activity; IEA.
DR   GO; GO:0016310; P:phosphorylation; IEA.
DR   InterPro; IPR007186; PMEI.
DR   InterPro; IPR006501; PME_inhib.
DR   InterPro; IPR002192; PPDK_N_term.
DR   Pfam; PF04043; PMEI; 1.
DR   Pfam; PF01326; PPDK_N; 1.
DR   TIGRFAMs; TIGR01614; PME_inhib; 1.
SQ   SEQUENCE   1540 AA;  172061 MW;  241B1433F0A2E79E CRC64;

  Query Match             57.1%;  Score 52;  DB 2;  Length 1540;
  Best Local Similarity   69.2%;  Pred. No. 15;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 RWEDPGKQLYNVE 15
              ||| |||:|| |
Db        357 RWERKGKQMYNPE 369


RESULT 15
Q9MXA7
ID   Q9MXA7      PRELIMINARY;      PRT;    92 AA.
AC   Q9MXA7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   MHC class I antigen (Fragment).
GN   Name=Bain-UA*L13;
OS   Barbus intermedius (Lake tana barbels).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Barbus.
OX   NCBI_TaxID=40831;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kruiswijk C.P., Stet R.J.M.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AJ007897; CAB97341.1; -.
DR   HSSP; P01897; 1LDP.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
FT   NON_TER       1      1
FT   NON_TER      92     92
SQ   SEQUENCE   92 AA;  10463 MW;  A1D08F3030F9E144 CRC64;

  Query Match             54.9%;  Score 50;  DB 2;  Length 92;
  Best Local Similarity   50.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||:|:  | |: ||:|
Db         53 KNKWDSTGAQINNVKA 68


Search completed: August 24, 2005, 23:43:28
Job time : 172 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 24, 2005, 23:36:09 ; Search time 42 Seconds
(without alignments)
28.438 Million cell updates/sec

Title: US-09-865-281A-1
Perfect score: 91
Sequence: 1 KNRWEDPGKQLYNVEA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	100.0	16	3	US-09-070-907-1	Sequence 1, Appli
2	91	100.0	63	1	US-08-447-411-24	Sequence 24, Appl
3	91	100.0	63	1	US-08-447-411-63	Sequence 63, Appl
4	91	100.0	63	2	US-08-662-227-20	Sequence 20, Appl
5	91	100.0	63	3	US-09-017-947-20	Sequence 20, Appl
6	91	100.0	63	4	US-09-925-442-20	Sequence 20, Appl
7	91	100.0	310	4	US-09-834-309-7	Sequence 7, Appli
8	91	100.0	310	4	US-09-834-309-8	Sequence 8, Appli
9	91	100.0	310	4	US-09-834-309-9	Sequence 9, Appli
10	91	100.0	1663	2	US-08-793-126-1	Sequence 1, Appli
11	91	100.0	1663	3	US-09-132-271-1	Sequence 1, Appli
12	91	100.0	1663	3	US-09-142-334-22	Sequence 22, Appl
13	80	87.9	63	1	US-08-447-411-26	Sequence 26, Appl
14	79	86.8	63	1	US-08-447-411-27	Sequence 27, Appl
15	73	80.2	63	1	US-08-447-411-25	Sequence 25, Appl
16	73	80.2	308	4	US-09-582-761B-26	Sequence 26, Appl
17	73	80.2	330	4	US-09-582-761B-37	Sequence 37, Appl
18	73	80.2	929	4	US-09-582-761B-27	Sequence 27, Appl
19	60	65.9	11	4	US-09-039-060A-6	Sequence 6, Appli
20	60	65.9	11	5	PCT-US94-01234-37	Sequence 37, Appl
21	60	65.9	11	5	PCT-US94-01263-7	Sequence 7, Appli
22	52	57.1	1333	1	US-08-447-411-76	Sequence 76, Appl
23	52	57.1	1333	2	US-08-662-227-34	Sequence 34, Appl
24	52	57.1	1333	3	US-09-017-947-34	Sequence 34, Appl
25	52	57.1	1333	4	US-09-925-442-34	Sequence 34, Appl
26	51	56.0	10	1	US-08-634-060-33	Sequence 33, Appl
27	51	56.0	10	2	US-08-700-846-5	Sequence 5, Appli
28	46	50.5	1493	4	US-09-713-273A-20	Sequence 20, Appl
29	44	48.4	193	4	US-09-248-796A-20794	Sequence 20794, A
30	44	48.4	2628	3	US-09-413-814-11	Sequence 11, Appl
31	43	47.3	28	2	US-08-448-603A-7	Sequence 7, Appli
32	43	47.3	28	3	US-09-134-075-7	Sequence 7, Appli
33	43	47.3	28	3	US-09-492-739-7	Sequence 7, Appli
34	43	47.3	28	4	US-09-966-931A-7	Sequence 7, Appli
35	43	47.3	63	1	US-08-447-411-23	Sequence 23, Appl
36	43	47.3	63	1	US-08-447-411-62	Sequence 62, Appl
37	43	47.3	63	2	US-08-662-227-19	Sequence 19, Appl
38	43	47.3	63	3	US-09-017-947-19	Sequence 19, Appl
39	43	47.3	63	4	US-09-925-442-19	Sequence 19, Appl
40	43	47.3	95	4	US-09-270-767-39763	Sequence 39763, A
41	43	47.3	95	4	US-09-270-767-54980	Sequence 54980, A
42	43	47.3	469	3	US-08-889-841B-23	Sequence 23, Appl
43	43	47.3	469	4	US-09-419-362-23	Sequence 23, Appl
44	43	47.3	494	3	US-08-889-841B-19	Sequence 19, Appl
45	43	47.3	494	4	US-09-419-362-19	Sequence 19, Appl

ALIGNMENTS

```
RESULT 1
US-09-070-907-1
; Sequence 1, Application US/09070907
; Patent No. 6238667
; GENERAL INFORMATION:
;  APPLICANT: Kohler, Heinz
;  TITLE OF INVENTION: METHOD OF AFFINITY CROSS-LINKING BIOLOGICALLY ACTIVE
;  TITLE OF INVENTION: IMMUNOGENIC PEPTIDES TO ANTIBODIES.
;  FILE REFERENCE: 35629
;  CURRENT APPLICATION NUMBER: US/09/070,907
;  CURRENT FILING DATE: 1998-05-04
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 1
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: AMINO ACID
;   OTHER INFORMATION: SEQUENCE DERIVED FROM Cd3 peptide
US-09-070-907-1

  Query Match             100.0%;  Score 91;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-09;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          1 KNRWEDPGKQLYNVEA 16


RESULT 2
US-08-447-411-24
; Sequence 24, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/447,411
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/043,747
;       FILING DATE:  07-APR-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Oblon, No. 5773243man F.
;       REGISTRATION NUMBER:  24,618
;       REFERENCE/DOCKET NUMBER:  1126-101-0
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 413-3000
;       TELEFAX:  (703) 413-2220
;       TELEX:  248855 OPAT UR
;   INFORMATION FOR SEQ ID NO:  24:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  63 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
US-08-447-411-24

  Query Match             100.0%;  Score 91;  DB 1;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-08;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          9 KNRWEDPGKQLYNVEA 24


RESULT 3
US-08-447-411-63
; Sequence 63, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  63 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-08-447-411-63

  Query Match             100.0%;  Score 91;  DB 1;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-08;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          9 KNRWEDPGKQLYNVEA 24


RESULT 4
US-08-662-227-20
; Sequence 20, Application US/08662227
; Patent No. 5922320
;  GENERAL INFORMATION:
;    APPLICANT:  VOGEL, CARL-WILHELM
;    APPLICANT:  BREDEHORST, REINHORST
;    APPLICANT:  KOCK, MICHAEL
;    APPLICANT:  FRITZINGER, DAVID
;    TITLE OF INVENTION:  RECOMBINANT PROCVF
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/662,227
;      FILING DATE:  14-JUN-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-0107-0X
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  63 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-662-227-20

  Query Match             100.0%;  Score 91;  DB 2;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-08;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          9 KNRWEDPGKQLYNVEA 24
```

```
RESULT 5
US-09-017-947-20
; Sequence 20, Application US/09017947
; Patent No. 6303754
;  GENERAL INFORMATION:
;    APPLICANT:  VOGEL, CARL-WILHELM
;    APPLICANT:  BREDEHORST, REINHORST
;    APPLICANT:  KOCK, MICHAEL
;    APPLICANT:  FRITZINGER, DAVID
;    TITLE OF INVENTION:  RECOMBINANT PROCVF
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/017,947
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/662,227
;      FILING DATE:  14-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-0107-0X
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  63 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-017-947-20

  Query Match             100.0%;  Score 91;  DB 3;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-08;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          9 KNRWEDPGKQLYNVEA 24


RESULT 6
US-09-925-442-20
; Sequence 20, Application US/09925442
; Patent No. 6607897
;   GENERAL INFORMATION:
;        APPLICANT: VOGEL, CARL-WILHELM
;                   BREDEHORST, REINHORST
;                   KOCK, MICHAEL
;                   FRITZINGER, DAVID
;        TITLE OF INVENTION: RECOMBINANT PROCVF
;        NUMBER OF SEQUENCES: 39
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;                        P.C.
;             STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
;             CITY: ARLINGTON
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22202
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/925,442
;             FILING DATE: 10-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/017,947
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: OBLON, NORMAN F.
;             REGISTRATION NUMBER: 24,618
;             REFERENCE/DOCKET NUMBER: 1126-0107-0X
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-413-3000
;             TELEFAX: 703-413-2220
;   INFORMATION FOR SEQ ID NO: 20:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 63 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-925-442-20

  Query Match             100.0%;  Score 91;  DB 4;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-08;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          9 KNRWEDPGKQLYNVEA 24


RESULT 7
US-09-834-309-7
; Sequence 7, Application US/09834309
; Patent No. 6820011
; GENERAL INFORMATION:
;  APPLICANT: Chen, Xiaojiang
;  APPLICANT:  Holers, V. Michael
;  TITLE OF INVENTION: THREE-DIMENSIONAL STRUCTURE OF COMPLEMENT RECEPTOR TYPE 2 AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 2848-43
;  CURRENT APPLICATION NUMBER: US/09/834,309
;  CURRENT FILING DATE:  2001-04-11
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 310
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-834-309-7

  Query Match             100.0%;  Score 91;  DB 4;  Length 310;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239


RESULT 8
US-09-834-309-8
```

```
; Sequence 8, Application US/09834309
; Patent No. 6820011
; GENERAL INFORMATION:
;  APPLICANT: Chen, Xiaojiang
;  APPLICANT:  Holers, V. Michael
;  TITLE OF INVENTION: THREE-DIMENSIONAL STRUCTURE OF COMPLEMENT RECEPTOR TYPE 2 AND USE
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 2848-43
;  CURRENT APPLICATION NUMBER: US/09/834,309
;  CURRENT FILING DATE:  2001-04-11
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
;   LENGTH: 310
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-834-309-8

  Query Match             100.0%;  Score 91;  DB 4;  Length 310;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239


RESULT 9
US-09-834-309-9
; Sequence 9, Application US/09834309
; Patent No. 6820011
; GENERAL INFORMATION:
;  APPLICANT: Chen, Xiaojiang
;  APPLICANT:  Holers, V. Michael
;  TITLE OF INVENTION: THREE-DIMENSIONAL STRUCTURE OF COMPLEMENT RECEPTOR TYPE 2 AND USE
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 2848-43
;  CURRENT APPLICATION NUMBER: US/09/834,309
;  CURRENT FILING DATE:  2001-04-11
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 310
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-834-309-9

  Query Match             100.0%;  Score 91;  DB 4;  Length 310;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239


RESULT 10
US-08-793-126-1
; Sequence 1, Application US/08793126
; Patent No. 5849297
;  GENERAL INFORMATION:
;    APPLICANT:  Harrison, Richard Alexander
;    APPLICANT:  Farries, Charles Timothy
;    TITLE OF INVENTION:  MODIFIED HUMAN C3 PROTEINS
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HALE AND DORR LLP
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  United States of America
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,126
;      FILING DATE:  07-FEB-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baker, Hollie L.
;      REGISTRATION NUMBER:  31,321
;      REFERENCE/DOCKET NUMBER:  102286.377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 526-6000
;      TELEFAX:  (617) 526-5000
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1663 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-793-126-1

  Query Match             100.0%;  Score 91;  DB 2;  Length 1663;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db       1217 KNRWEDPGKQLYNVEA 1232


RESULT 11
US-09-132-271-1
; Sequence 1, Application US/09132271
; Patent No. 6221657
;  GENERAL INFORMATION:
;    APPLICANT:  Harrison, Richard Alexander
;    APPLICANT:  Farries, Charles Timothy
;    TITLE OF INVENTION:  MODIFIED HUMAN C3 PROTEINS
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HALE AND DORR LLP
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  United States of America
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/132,271
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,126
;      FILING DATE:  07-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baker, Hollie L.
;      REGISTRATION NUMBER:  31,321
;      REFERENCE/DOCKET NUMBER:  102286.377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 526-6000
;      TELEFAX:  (617) 526-5000
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1663 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-132-271-1

  Query Match             100.0%;  Score 91;  DB 3;  Length 1663;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db       1217 KNRWEDPGKQLYNVEA 1232


RESULT 12
US-09-142-334-22
; Sequence 22, Application US/09142334
; Patent No. 6268485
; GENERAL INFORMATION:
;  APPLICANT: Farries, Timothy C.
;  APPLICANT: Harrison, Richard A.
;  TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
;  FILE REFERENCE: 4-30443/A/IMU/PCT
;  CURRENT APPLICATION NUMBER: US/09/142,334
;  CURRENT FILING DATE: 1999-04-15
;  EARLIER APPLICATION NUMBER: PCT/GB97/00603
;  EARLIER FILING DATE: 1997-03-04
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 1663
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-142-334-22

  Query Match             100.0%;  Score 91;  DB 3;  Length 1663;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db       1217 KNRWEDPGKQLYNVEA 1232


RESULT 13
US-08-447-411-26
; Sequence 26, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  63 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-447-411-26

  Query Match             87.9%;  Score 80;  DB 1;  Length 63;
  Best Local Similarity   81.2%;  Pred. No. 2.4e-06;
  Matches   13;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              :|||:||:|||||||
Db          9 RNRWEEPGQQLYNVEA 24


RESULT 14
US-08-447-411-27
; Sequence 27, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  63 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Oryctolagus cuniculus
US-08-447-411-27
```

```
  Query Match             86.8%;  Score 79;  DB 1;  Length 63;
  Best Local Similarity   81.2%;  Pred. No. 3.5e-06;
  Matches   13;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              |||||:||::||||||
Db          9 KNRWEEPGQRLYNVEA 24


RESULT 15
US-08-447-411-25
; Sequence 25, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  63 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-447-411-25

  Query Match             80.2%;  Score 73;  DB 1;  Length 63;
  Best Local Similarity   75.0%;  Pred. No. 3.7e-05;
  Matches   12;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              :||||:| :|||||||
Db          9 RNRWEEPDQQLYNVEA 24


Search completed: August 24, 2005, 23:44:59
Job time : 43 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 24, 2005, 23:43:35 ; Search time 159 Seconds
(without alignments)
39.405 Million cell updates/sec

Title: US-09-865-281A-1
Perfect score: 91
Sequence: 1 KNRWEDPGKQLYNVEA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	100.0	16	10	US-09-865-281A-1	Sequence 1, Appli
2	91	100.0	16	17	US-10-795-081A-1	Sequence 1, Appli
3	91	100.0	63	9	US-09-925-442-20	Sequence 20, Appl
4	91	100.0	94	15	US-10-424-599-219407	Sequence 219407,
5	91	100.0	310	11	US-09-834-309-7	Sequence 7, Appli
6	91	100.0	310	11	US-09-834-309-8	Sequence 8, Appli
7	91	100.0	310	11	US-09-834-309-9	Sequence 9, Appli
8	91	100.0	705	15	US-10-379-747-4	Sequence 4, Appli
9	91	100.0	935	18	US-10-887-775-32	Sequence 32, Appl
10	91	100.0	1255	17	US-10-497-073-17	Sequence 17, Appl
11	91	100.0	1288	17	US-10-741-600-1326	Sequence 1326, Ap
12	91	100.0	1638	17	US-10-884-813-8	Sequence 8, Appli
13	91	100.0	1638	17	US-10-884-813-12	Sequence 12, Appl
14	91	100.0	1663	9	US-09-875-519A-22	Sequence 22, Appl
15	91	100.0	1663	10	US-09-842-758-41	Sequence 41, Appl
16	91	100.0	1663	15	US-10-379-747-2	Sequence 2, Appli
17	91	100.0	1663	15	US-10-174-333-41	Sequence 41, Appl
18	91	100.0	1663	17	US-10-741-600-1327	Sequence 1327, Ap
19	91	100.0	1663	17	US-10-928-312-2	Sequence 2, Appli
20	91	100.0	1663	17	US-10-884-813-2	Sequence 2, Appli
21	91	100.0	1663	17	US-10-884-813-6	Sequence 6, Appli
22	91	100.0	1663	17	US-10-884-813-10	Sequence 10, Appl
23	91	100.0	1663	18	US-10-887-775-30	Sequence 30, Appl
24	83	91.2	296	15	US-10-398-916-29	Sequence 29, Appl
25	83	91.2	296	15	US-10-398-916-30	Sequence 30, Appl
26	83	91.2	300	15	US-10-398-916-13	Sequence 13, Appl
27	82	90.1	105	9	US-09-925-301-1490	Sequence 1490, Ap
28	76	83.5	300	15	US-10-398-916-11	Sequence 11, Appl
29	76	83.5	300	15	US-10-398-916-15	Sequence 15, Appl
30	73	80.2	300	15	US-10-398-916-9	Sequence 9, Appli
31	73	80.2	312	15	US-10-398-916-17	Sequence 17, Appl
32	73	80.2	409	16	US-10-466-655-6	Sequence 6, Appli
33	73	80.2	1663	10	US-09-842-758-43	Sequence 43, Appl
34	73	80.2	1663	15	US-10-174-333-43	Sequence 43, Appl
35	71	78.0	1661	10	US-09-842-758-42	Sequence 42, Appl
36	71	78.0	1661	15	US-10-174-333-42	Sequence 42, Appl
37	69	75.8	296	18	US-10-505-546-10	Sequence 10, Appl
38	60	65.9	11	15	US-10-408-849-6	Sequence 6, Appli
39	52	57.1	1333	9	US-09-925-442-34	Sequence 34, Appl
40	46.5	51.1	500	15	US-10-282-122A-57243	Sequence 57243, A
41	46	50.5	329	16	US-10-425-115-197051	Sequence 197051,
42	46	50.5	437	15	US-10-424-599-190068	Sequence 190068,
43	46	50.5	1493	15	US-10-607-095-20	Sequence 20, Appl
44	44	48.4	91	11	US-09-864-408A-3792	Sequence 3792, Ap
45	43	47.3	28	10	US-09-966-931-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-865-281A-1
; Sequence 1, Application US/09865281A
; Publication No. US20030103984A1
; GENERAL INFORMATION:
; APPLICANT: Kohler, Heinz
; TITLE OF INVENTION: FUSION PROTEINS OF BIOLOGICALLY ACTIVE PEPTIDES AND ANTIBODIES
; FILE REFERENCE: 411.35629PC2
; CURRENT APPLICATION NUMBER: US/09/865,281A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 09/070,907
; PRIOR FILING DATE: 1998-05-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Synthesized peptide with sequence derived from position 1217-1232
US-09-865-281A-1

Query Match 100.0%; Score 91; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 1 KNRWEDPGKQLYNVEA 16

RESULT 2

```
US-10-795-081A-1
; Sequence 1, Application US/10795081A
; Publication No. US20050033033A1
; GENERAL INFORMATION:
;  APPLICANT: Kohler, Heinz
;  TITLE OF INVENTION: TRANS-MEMBRANE-ANTIBODY INDUCED INHIBITION OF APOPTOSIS
;  FILE REFERENCE: 411.35629AP3
;  CURRENT APPLICATION NUMBER: US/10/795,081A
;  CURRENT FILING DATE:  2004-03-05
;  PRIOR APPLICATION NUMBER: 60/451,980
;  PRIOR FILING DATE: 2003-03-05
;  PRIOR APPLICATION NUMBER: 09/865,281
;  PRIOR FILING DATE: 2001-05-29
;  PRIOR APPLICATION NUMBER: 09/070,907
;  PRIOR FILING DATE: 1998-05-04
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: Patentln version 3.0
; SEQ ID NO 1
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Synthesized peptide with sequence derived from position 1217-1232
US-10-795-081A-1

  Query Match             100.0%;  Score 91;  DB 17;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          1 KNRWEDPGKQLYNVEA 16


RESULT 3
US-09-925-442-20
; Sequence 20, Application US/09925442
; Patent No. US20020103346A1
;   GENERAL INFORMATION:
;        APPLICANT: VOGEL, CARL-WILHELM
;                   BREDEHORST, REINHORST
;                   KOCK, MICHAEL
;                   FRITZINGER, DAVID
;        TITLE OF INVENTION: RECOMBINANT PROCVF
;        NUMBER OF SEQUENCES: 39
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;                        P.C.
;             STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
;             CITY: ARLINGTON
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22202
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/925,442
;             FILING DATE: 10-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/017,947
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: OBLON, NORMAN F.
;             REGISTRATION NUMBER: 24,618
;             REFERENCE/DOCKET NUMBER: 1126-0107-0X
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-413-3000
;             TELEFAX: 703-413-2220
;   INFORMATION FOR SEQ ID NO: 20:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 63 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-925-442-20

  Query Match             100.0%;  Score 91;  DB 9;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          9 KNRWEDPGKQLYNVEA 24


RESULT 4
US-10-424-599-219407
; Sequence 219407, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 219407
;   LENGTH: 94
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(94)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_40150C.1.pep
US-10-424-599-219407

  Query Match             100.0%;  Score 91;  DB 15;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 1e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db         43 KNRWEDPGKQLYNVEA 58


RESULT 5
US-09-834-309-7
; Sequence 7, Application US/09834309
; Publication No. US20040005538A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Xiaojiang
;  APPLICANT:  Holers, V. Michael
;  TITLE OF INVENTION: THREE-DIMENSIONAL STRUCTURE OF COMPLEMENT RECEPTOR TYPE 2 AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 2848-43
;  CURRENT APPLICATION NUMBER: US/09/834,309
;  CURRENT FILING DATE:  2001-04-11
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 310
;   TYPE: PRT
```

```
;   ORGANISM: Homo sapiens
US-09-834-309-7

  Query Match             100.0%;  Score 91;  DB 11;  Length 310;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239


RESULT 6
US-09-834-309-8
; Sequence 8, Application US/09834309
; Publication No. US20040005538A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Xiaojiang
;  APPLICANT:  Holers, V. Michael
;  TITLE OF INVENTION: THREE-DIMENSIONAL STRUCTURE OF COMPLEMENT RECEPTOR TYPE 2 AND USE
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 2848-43
;  CURRENT APPLICATION NUMBER: US/09/834,309
;  CURRENT FILING DATE:  2001-04-11
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
;   LENGTH: 310
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-834-309-8

  Query Match             100.0%;  Score 91;  DB 11;  Length 310;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239


RESULT 7
US-09-834-309-9
; Sequence 9, Application US/09834309
; Publication No. US20040005538A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Xiaojiang
;  APPLICANT:  Holers, V. Michael
;  TITLE OF INVENTION: THREE-DIMENSIONAL STRUCTURE OF COMPLEMENT RECEPTOR TYPE 2 AND USE
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 2848-43
;  CURRENT APPLICATION NUMBER: US/09/834,309
;  CURRENT FILING DATE:  2001-04-11
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 310
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-834-309-9

  Query Match             100.0%;  Score 91;  DB 11;  Length 310;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        224 KNRWEDPGKQLYNVEA 239


RESULT 8
US-10-379-747-4
; Sequence 4, Application US/10379747
; Publication No. US20040023874A1
; GENERAL INFORMATION:
;  APPLICANT: Burgess, Catherine E.;
;  APPLICANT:  Chant, John S.;
;  APPLICANT:  Chaudhuri, Amitabha ;
;  APPLICANT:  Edinger, Shlomit R.;
;  APPLICANT:  Gangolli, Esha A.;
;  APPLICANT:  Malyankar, Uriel M.;
;  APPLICANT:  Miller, Charles E.;
;  APPLICANT:  Ooi, Chean Eng;
;  APPLICANT:  Ort, Tatiana A.;
;  APPLICANT:  Patturajan, Meera ;
;  APPLICANT:  Rastelli, Luca ;
;  APPLICANT:  Rieger, Daniel K.;
;  APPLICANT:  Shimkets, Richard A.;
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
;  FILE REFERENCE: 21402-568B
;  CURRENT APPLICATION NUMBER: US/10/379,747
;  CURRENT FILING DATE:  2003-03-05
;  PRIOR APPLICATION NUMBER: 60/365,034
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: 60/366,420
;  PRIOR FILING DATE: 2002-03-21
;  PRIOR APPLICATION NUMBER: 60/365,477
;  PRIOR FILING DATE: 2002-03-19
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 4
;   LENGTH: 705
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-379-747-4

  Query Match             100.0%;  Score 91;  DB 15;  Length 705;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        259 KNRWEDPGKQLYNVEA 274


RESULT 9
US-10-887-775-32
; Sequence 32, Application US/10887775
; Publication No. US20050130182A1
; GENERAL INFORMATION:
;  APPLICANT: MESSER, Jeffrey
;  APPLICANT:  BENJAMIN, Dennis
;  APPLICANT:  VATH, James
;  APPLICANT:  SIGEL, Eric
;  TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  ENDOMETRIOSIS
;  FILE REFERENCE: PPI-149
;  CURRENT APPLICATION NUMBER: US/10/887,775
;  CURRENT FILING DATE:  2004-07-09
;  PRIOR APPLICATION NUMBER: 60/486,379
;  PRIOR FILING DATE: 2003-07-11
;  PRIOR APPLICATION NUMBER: 60/533,430
;  PRIOR FILING DATE: 2003-12-29
;  PRIOR APPLICATION NUMBER: 60/575,269
;  PRIOR FILING DATE: 2004-05-08
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32
;   LENGTH: 935
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-887-775-32

  Query Match             100.0%;  Score 91;  DB 18;  Length 935;
```

```
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        489 KNRWEDPGKQLYNVEA 504


RESULT 10
US-10-497-073-17
; Sequence 17, Application US/10497073
; Publication No. US20050048584A1
; GENERAL INFORMATION:
;  APPLICANT: BioVisioN AG
;  TITLE OF INVENTION: Method for detecting Alzheimer's disease and differentiating
;  TITLE OF INVENTION:  Alzheimer's disease from other demential diseases, associated
;  TITLE OF INVENTION:  peptides and the use thereof
;  FILE REFERENCE: C3f-PCT
;  CURRENT APPLICATION NUMBER: US/10/497,073
;  CURRENT FILING DATE:  2004-05-28
;  PRIOR APPLICATION NUMBER: DE10158180
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: PCT/DE02/04360
;  PRIOR FILING DATE: 2002-11-27
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-497-073-17

  Query Match             100.0%;  Score 91;  DB 17;  Length 1255;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db        809 KNRWEDPGKQLYNVEA 824


RESULT 11
US-10-741-600-1326
; Sequence 1326, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1326
;   LENGTH: 1288
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-741-600-1326

  Query Match             100.0%;  Score 91;  DB 17;  Length 1288;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db       1217 KNRWEDPGKQLYNVEA 1232


RESULT 12
US-10-884-813-8
; Sequence 8, Application US/10884813
; Publication No. US20050079585A1
; GENERAL INFORMATION:
;  APPLICANT: Kolln, Johanna
;  APPLICANT:  Bredehorst, Reinhard
;  APPLICANT:  Spillner, Edzard
;  TITLE OF INVENTION: Complement Depletion with Recombinant Human C3 Derivatives
;  FILE REFERENCE: P 63782
;  CURRENT APPLICATION NUMBER: US/10/884,813
;  CURRENT FILING DATE:  2004-07-02
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 1638
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hybrid protein
US-10-884-813-8

  Query Match             100.0%;  Score 91;  DB 17;  Length 1638;
  Best Local Similarity   100.0%;  Pred. No. 2e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db       1217 KNRWEDPGKQLYNVEA 1232


RESULT 13
US-10-884-813-12
; Sequence 12, Application US/10884813
; Publication No. US20050079585A1
; GENERAL INFORMATION:
;  APPLICANT: Kolln, Johanna
;  APPLICANT:  Bredehorst, Reinhard
;  APPLICANT:  Spillner, Edzard
;  TITLE OF INVENTION: Complement Depletion with Recombinant Human C3 Derivatives
;  FILE REFERENCE: P 63782
;  CURRENT APPLICATION NUMBER: US/10/884,813
;  CURRENT FILING DATE:  2004-07-02
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 1638
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hybrid protein
US-10-884-813-12

  Query Match             100.0%;  Score 91;  DB 17;  Length 1638;
  Best Local Similarity   100.0%;  Pred. No. 2e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db       1217 KNRWEDPGKQLYNVEA 1232


RESULT 14
US-09-875-519A-22
; Sequence 22, Application US/09875519A
; Patent No. US20020068059A1
; GENERAL INFORMATION:
;  APPLICANT: Farries, Timothy C.
;  APPLICANT:  Harrison, Richard A.
;  TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
;  FILE REFERENCE: 4-30443/A/IMU/PCT
;  CURRENT APPLICATION NUMBER: US/09/875,519A
;  CURRENT FILING DATE:  2001-06-06
;  PRIOR APPLICATION NUMBER: PCT/GB97/00603
;  PRIOR FILING DATE: 1997-03-04
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 22
;   LENGTH: 1663
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-875-519A-22

  Query Match             100.0%;  Score 91;  DB 9;  Length 1663;
  Best Local Similarity   100.0%;  Pred. No. 2e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db       1217 KNRWEDPGKQLYNVEA 1232


RESULT 15
US-09-842-758-41
; Sequence 41, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
;  APPLICANT: Vernet, Corine A. M.
;  APPLICANT:  Fernandes, Elma R
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Boldog, Ferenc L
;  APPLICANT:  Zerhusen, Bryan D
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Majumder, Kumud
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Gangolli, Esha A
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Taupier, Raymond J
;  APPLICANT:  Grosse, William M
;  APPLICANT:  Edward, Szekeres S
;  APPLICANT:  Alsobrook II, John P
;  TITLE OF INVENTION: No. US20030083244A1el Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-783
;  CURRENT APPLICATION NUMBER: US/09/842,758
;  CURRENT FILING DATE:  2001-04-25
;  PRIOR APPLICATION NUMBER: 60/200,158
;  PRIOR FILING DATE: 2000-04-26
;  PRIOR APPLICATION NUMBER: 60/200,613
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,780
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/201,006
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/201,007
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/201,236
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/201,238
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/201,186
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: 60/201,474
;  PRIOR FILING DATE: 2000-05-03
;  PRIOR APPLICATION NUMBER: 60/201,508
;  PRIOR FILING DATE: 2000-05-03
;  PRIOR APPLICATION NUMBER: 60/220,591
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: 60/232,678
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: 60/263,217
;  PRIOR FILING DATE: 2001-01-22
;  PRIOR APPLICATION NUMBER: 60/265,160
;  PRIOR FILING DATE: 2001-01-30
;  NUMBER OF SEQ ID NOS: 113
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
;   LENGTH: 1663
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-842-758-41

  Query Match             100.0%;  Score 91;  DB 10;  Length 1663;
  Best Local Similarity   100.0%;  Pred. No. 2e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db       1217 KNRWEDPGKQLYNVEA 1232


Search completed: August 24, 2005, 23:57:19
Job time : 160 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 24, 2005, 23:44:21 ; Search time 161 Seconds
(without alignments)
38.436 Million cell updates/sec

Title: US-09-865-281A-1
Perfect score: 91
Sequence: 1 KNRWEDPGKQLYNVEA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 649094

Minimum DB seq length: 0
Maximum DB seq length: 16

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	100.0	16	4	AAB92360	Aab92360 Miscellan
2	91	100.0	16	6	ABP58217	Abp58217 Immunosti
3	91	100.0	16	8	ADS17594	Ads17594 Peptide d
4	64	70.3	12	5	AAU74853	Aau74853 Complemen
5	60	65.9	11	2	AAR57873	Aar57873 CR2 cell
6	60	65.9	11	2	AAR57904	Aar57904 CR2 recep
7	60	65.9	11	8	ADH73668	Adh73668 Novel rec
8	55	60.4	11	2	AAR95868	Aar95868 CR2 recep
9	55	60.4	11	2	AAW27141	Aaw27141 Complemen
10	55	60.4	11	2	AAW87720	Aaw87720 Epitope i
11	51	56.0	10	2	AAW46335	Aaw46335 Binding d
12	46	50.5	14	2	AAR95584	Aar95584 PepC3 der
13	38	41.8	16	2	AAW32826	Aaw32826 HIV-1 CDC
14	37	40.7	11	4	ABP18544	Abp18544 HIV B62 s
15	37	40.7	15	2	AAR24423	Aar24423 Sequence
16	37	40.7	15	2	AAR32415	Aar32415 Sequence
17	37	40.7	15	2	AAR32399	Aar32399 Sequence
18	37	40.7	15	2	AAW76983	Aaw76983 Fusion im
19	37	40.7	15	2	AAW76981	Aaw76981 Fusion im
20	37	40.7	15	3	AAY66444	Aay66444 HLA-A2-bi
21	37	40.7	15	4	ABP24898	Abp24898 HIV DR 3a
22	37	40.7	16	1	AAP82479	Aap82479 Peptide c
23	37	40.7	16	2	AAR24424	Aar24424 Sequence
24	37	40.7	16	2	AAR85369	Aar85369 HTLV-IIIB
25	37	40.7	16	2	AAW07391	Aaw07391 HIV-1 CD4
26	37	40.7	16	2	AAW10345	Aaw10345 HIV epito
27	37	40.7	16	2	AAW16512	Aaw16512 HTLV-IIIB
28	37	40.7	16	2	AAW32824	Aaw32824 HIV-1 SC
29	37	40.7	16	2	AAW16535	Aaw16535 HIV-1 BH1
30	37	40.7	16	2	AAW32825	Aaw32825 HIV-1 SF2
31	37	40.7	16	2	AAW32822	Aaw32822 HIV-1 BRU
32	37	40.7	16	2	AAW32823	Aaw32823 HIV-1 MN
33	37	40.7	16	2	AAW32828	Aaw32828 HIV-1 RF
34	37	40.7	16	2	AAW53140	Aaw53140 HIV gp160
35	37	40.7	16	2	AAW85381	Aaw85381 Helper T-
36	37	40.7	16	2	AAW76982	Aaw76982 Fusion im
37	37	40.7	16	2	AAW54937	Aaw54937 HIV gp120
38	37	40.7	16	2	AAY04046	Aay04046 Covalentl
39	37	40.7	16	3	AAY73159	Aay73159 HIV-deriv
40	37	40.7	16	4	AAB49073	Aab49073 HIV gp120
41	37	40.7	16	4	AAB46174	Aab46174 HIV gp120
42	37	40.7	16	4	AAU12518	Aau12518 Human HIV
43	37	40.7	16	4	AAU12526	Aau12526 Human HIV
44	37	40.7	16	4	AAU12495	Aau12495 Human HIV
45	37	40.7	16	4	AAU12540	Aau12540 Human HIV

ALIGNMENTS

RESULT 1
AAB92360
ID AAB92360 standard; peptide; 16 AA.
XX
AC AAB92360;
XX
DT 22-JUN-2001 (first entry)
XX
DE Miscellaneous peptide SEQ ID NO:1536.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US013576.
XX
PR 17-MAY-1999; 99US-0134406P.
PR 10-SEP-1999; 99US-0153406P.
PR 15-OCT-1999; 99US-0159783P.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity.
XX
PS Disclosure; Page 707; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity in
CC vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent

CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 91; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 1 KNRWEDPGKQLYNVEA 16

RESULT 2
ABP58217
ID ABP58217 standard; peptide; 16 AA.
XX
AC ABP58217;
XX
DT 21-MAR-2003 (first entry)
XX
DE Immunostimulant C3d peptide.
XX
KW Immunostimulant; C3d; human; fusion protein; tumour; vaccine; adjuvant.
XX
OS Homo sapiens.
XX
PN WO200297041-A2.
XX
PD 05-DEC-2002.
XX
PF 29-MAY-2002; 2002WO-US016651.
XX
PR 29-MAY-2001; 2001US-00865281.
XX
PA (IMMP-) IMMPHERON INC.
PA (INNE-) INNEXUS CORP.
XX
PI Kohler H, Morgan C;
XX
DR WPI; 2003-140458/13.
XX
PT Novel fusion protein for use as molecular adjuvant, has an antibody and a
PT peptide with immunostimulatory, membrane transport or homophilic
PT activities, connected to the antibody by peptide bonds.
XX
PS Example 1; Page 14; 39pp; English.
XX
CC The present invention provides a fusion protein made up of an antibody
CC and a peptide having e.g. immunostimulant, membrane transport or
CC homophilic activity. The peptide is located at a site in the antibody
CC such that it does not compromise the antigen recognition of the antibody.
CC In order to enhance its activity, the peptide may be flanked by loop-
CC forming or conformation-conferring sequences. The present sequence is an
CC example of a suitable imunostimulatory peptide for use as a fusion
CC protein component. The peptide is derived from human C3d amino acids 1217
CC -1232. In examples from the invention, the C3d peptide was affinity cross
CC -linked to tumour anti-idiotype and tumour idiotype vaccine antibodies,
CC significantly enhancing the immune response to the tumour and protecting
CC against tumour challenge. The vaccination protocol did not include any
CC adjuvant, such as Freund's adjuvant or keyhole limpet haemocyanin
CC conjugation, both of which are not permissible by the FDA for human use
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 91; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 1 KNRWEDPGKQLYNVEA 16

RESULT 3
ADS17594
ID ADS17594 standard; peptide; 16 AA.
XX
AC ADS17594;
XX
DT 02-DEC-2004 (first entry)
XX
DE Peptide derived from the C3d peptide and affinity linked to 3H1 antibody.
XX
KW immunostimulatory; membrane transport; homophilic; signaling protein;
KW caspase; kinase; phosphatase; viral protein; tumour antigen;
KW nuclear protein; nucleolar protein; DNA synthesis; cytoskeletal protein;
KW cell proliferation; cytostasis; membrane transporter peptide;
KW Kaposi fibroblast factor; TAT peptide; HIV-1; antennapedia homeodomain;
KW herpes virus protein VP22; transportan peptide; Alzheimer's disease;
KW Huntington's disease; Parkinson's disease; C3d; 3H1; monoclonal antibody;
KW anti-idiotypic antibody; carcino-embryonic antigen; CEA;
KW anti-idiotype vaccine; antibody.
XX
OS Synthetic.
XX
PN WO2004078146-A2.
XX
PD 16-SEP-2004.
XX
PF 05-MAR-2004; 2004WO-US006911.
XX
PR 05-MAR-2003; 2003US-0451980P.
XX
PA (INNE-) INNEXUS BIOTECHNOLOGY INC.
PA (IMMP-) IMMPHERON INC.
XX
PI Kohler H, Muller S, Brown TL, Zhao Y, Morgan AC;
XX
DR WPI; 2004-653567/63.
XX
PT New compound for regulating normal or infected cell function comprising
PT an antibody conjugated to a membrane transporter peptide, useful in
PT preparing a composition for treating or preventing human diseases, e.g.
PT Alzheimer's disease.
XX
PS Example 1; SEQ ID NO 1; 50pp; English.
XX
CC The specification describes a fusion protein for regulating normal or
CC infected cell function, comprising an antibody conjugated to a peptide
CC having immunostimulatory, membrane transport, and homophilic activities.
CC The antibody is immunospecific for a signaling protein internal cell
CC consisting of caspases, kinases or phosphatases, an immature viral
CC protein, a cell-surface or intracellular tumour antigen, a nuclear or
CC nucleolar protein participating in regulation of DNA synthesis and gene
CC expression, or a cytoskeletal protein participating in cell proliferation
CC or cytostasis. The peptide portion of the fusion protein is preferably a
CC membrane transporter peptide that is endogenous to Kaposi fibroblast
CC factor, TAT peptides of HIV-1, antennapedia homeodomain-derived peptide,
CC herpes virus protein VP22, or transportan peptide. Fusion protein of the
CC invention are useful for preparing a composition for treating or
CC preventing human diseases, e.g., Alzheimer's disease, Huntington's
CC disease or Parkinson's disease. The present sequence represents a peptide
CC derived from the C3d region 1217-1232, which was affinity cross-linked
CC with 3H1 monoclonal antibody to produce fusion proteins of the invention.
CC 3H1 is a murine anti-idiotypic antibody which mimics the carcino-
CC embryonic antigen (CEA), and induces anti-CEA antibodies. The resulting
CC Cd3-3H1 fusion protein was used to enhance an anti-idiotype vaccine.
XX
SQ Sequence 16 AA;

```
  Query Match            100.0%;  Score 91;  DB 8;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 5.6e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLYNVEA 16
              ||||||||||||||||
Db          1 KNRWEDPGKQLYNVEA 16


RESULT 4
AAU74853
ID   AAU74853 standard; peptide; 12 AA.
XX
AC   AAU74853;
XX
DT   09-APR-2002  (first entry)
XX
DE   Complement receptor 2 (CD21/CD2) associated, C3d peptide.
XX
KW   Complement; receptor; CD21; CD2; C3d; immune response; B cell stimulator;
KW   vaccine; CD21/CD19 complex; tumour; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200192295-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-CA000785.
XX
PR   30-MAY-2000; 2000US-0207434P.
XX
PA   (UTOR ) UNIV TORONTO.
XX
PI   Isenman DE,  Clemenza L;
XX
DR   WPI; 2002-114323/15.
XX
PT   Ligand useful for modulating immune response such as in the preparation
PT   of vaccine comprises CD21 contacting amino acid residues of C3d molecule.
XX
PS   Disclosure; Page 5; 53pp; English.
XX
CC   The invention describes a ligand of the complement receptor 2 (CD21 or
CC   CD2) comprising amino acid residues 36-39 and 160-167 of the C3d
CC   molecule. The ligand is useful in the manufacture of a medicament such as
CC   a vaccine for modulating the immune response of a host (preferably tumour
CC   vaccine), and as antigens in immunogenic compositions, therapeutics
CC   diagnostic reagents, in the generation of diagnostic agents and as cancer
CC   therapeutics. The ligand has the ability to bind CD21 and stimulate B
CC   cells through the CD21/CD19 complex. Non-naturally occurring ligands and
CC   site specific mutated analogues of C3d demonstrate an enhanced binding
CC   affinity for CD21 as compared to the binding affinity of a wild-type C3d
CC   molecule. The ligand alters the immunogenicity of an antigen, e.g. by
CC   inducing or enhancing an immune response to an antigen in a host and thus
CC   protects the host against disease caused by the pathogen. This sequence
CC   represents a peptide segment of C3d, a protein of the complement pathway,
CC   found to have a major role in the interaction of C3d with complement
CC   receptor 2 (CD21 or CD2), described in the method of the invention
XX
SQ   Sequence 12 AA;

  Query Match            70.3%;  Score 64;  DB 5;  Length 12;
  Best Local Similarity  100.0%;  Pred. No. 0.0061;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 EDPGKQLYNVEA 16
              ||||||||||||
Db          1 EDPGKQLYNVEA 12


RESULT 5
AAR57873
ID   AAR57873 standard; peptide; 11 AA.
XX
AC   AAR57873;
XX
DT   25-MAR-2003  (revised)
DT   28-MAR-1995  (first entry)
XX
DE   CR2 cell receptor minimum binding site #2 for EBV gp350/220.
XX
KW   Binding site; CDR; complementarity determining region; immunoglobulin;
KW   heavy; light; primer extension; PCR; amplify; fibronectin; vitronectin;
KW   RGD-dependant; integrin ligand; von Willebrand factor; EBV; gp350/220;
KW   envelope gylcoprotein; HIV; gp120; reovirus; hemagglutinin; insulin;
KW   cellular receptor; CR2; CD4; hormone; thyroid stimulating hormone; TSH;
KW   transferrin; apolipoprotein; apo E; apo AI; MHC; class I; class II;
KW   non-RGD-dependent; vitronectin receptor; alpha-v, beta-3; modulation;
KW   anti-gpIIb/IIIa; monoclonal antibody; MAb; platelet adhesion; cancer;
KW   coagulation; inflammation; anti-vitronectin; tumour cell adhesion;
KW   migration.
XX
OS   Homo sapiens.
XX
PN   WO9418221-A1.
XX
PD   18-AUG-1994.
XX
PF   02-FEB-1994;   94WO-US001258.
XX
PR   02-FEB-1993;   93US-00012566.
PR   28-JUN-1993;   93US-00084542.
XX
PA   (SCRI ) SCRIPPS RES INST.
XX
PI   Barbas CF,  Lerner RA;
XX
DR   WPI; 1994-279675/34.
XX
PT   Production of binding sites within CDR regions of immunoglobulins -
PT   displayed on the surface of filamentous phage particles, for inhibiting
PT   platelet aggregation and vitronectin binding.
XX
PS   Disclosure; Page 26; 207pp; English.
XX
CC   The sequences given in AAR57837-84 are binding sites which were used in
CC   the method of the invention for producing a polypeptide having a binding
CC   site capable of binding a preselected agent. Nucleotide sequences
CC   encoding these binding site peptides were introduced into a CDR region of
CC   a nucleic acid encoding an immunoglobulin heavy (H) or light (L) chain,
CC   by amplifying the CDR region by primer extension. Preferred binding sites
CC   are derived from the RGD-dependant integrin ligands, eg. fibronectin,
CC   vitronectin, von Willebrand factor, from the envelope gylcoprotein from
CC   viruses such as HIV gp120, EBV gp350/ 220, reovirus hemagglutinin, from
CC   cellular receptors such as CR2 or CD4, from protein hormones such as
CC   thyroid stimulating hormone (TSH), insulin, transferrin, from
CC   apolipoproteins such as apo E and apo AI, from immunoglobulin CDRs and
CC   from MHC class I or II proteins. Non-RGD- dependent integrin binding
CC   sites were selected for the affinity to bind vitronectin receptor alpha-
CC   v, beta-3. An anti-gpIIb/IIIa monoclonal antibody (MAb) produced in this
CC   way can be used to modulate platelet adhesion in the treatment of
CC   coagulation and some inflammatory responses. An anti-vitronectin MAb can
CC   be used in the treatment of cancer by blocking tumour cell adhesion and
CC   migration. This sequence represents a binding site which mimics a binding
CC   site on the cell receptor CR2 which has binding specificity for the EBV
CC   gp350/220 receptor. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 11 AA;

  Query Match            65.9%;  Score 60;  DB 2;  Length 11;
  Best Local Similarity  100.0%;  Pred. No. 0.023;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 EDPGKQLYNVE 15
```

```
              ||||||||||
Db          1 EDPGKQLYNVE 11


RESULT 6
AAR57904
ID   AAR57904 standard; protein; 11 AA.
XX
AC   AAR57904;
XX
DT   16-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   30-MAR-1995  (first entry)
XX
DE   CR2 receptor-targeting peptide.
XX
KW   Adeno virus-2; Ad2; penton; receptor binding; epithelium; DNA delivery;
KW   gene transfer; gene therapy; antisense; antiviral therapy; CR2 receptor.
XX
OS   Human adenovirus type 2.
XX
PN   WO9417832-A1.
XX
PD   18-AUG-1994.
XX
PF   03-FEB-1994;   94WO-US001263.
XX
PR   09-FEB-1993;   93US-00015225.
PR   13-APR-1993;   93US-00046159.
XX
PA   (SCRI ) SCRIPPS RES INST.
XX
PI   Nemerow GR,  Wickham TJ;
XX
DR   WPI; 1994-279398/34.
XX
PT   Delivery of nucleotide sequences to mammalian cells - using a compsn
PT   comprising an adenovirus-derived protein and the nucleotide sequence.
XX
PS   Disclosure; Page 112; 111pp; English.
XX
CC   A coat protein subunit of Ad2, the penton, duplicates the epithelial cell
CC   receptor binding and DNA delivery properties of intact Ad2 virion and
CC   represents an improved means for gene therapy and antisense-based
CC   antiviral therapy. Compositions designed to target non-epithelial cells
CC   may include an Ad2-derived protein ligand conjugate. Polypeptides that
CC   include the sequences given in AAR57903-04 are capable of targeting CR2
CC   receptors and are useful in such compositions. (Updated on 25-MAR-2003 to
CC   correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             65.9%;  Score 60;  DB 2;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 EDPGKQLYNVE 15
              ||||||||||
Db          1 EDPGKQLYNVE 11


RESULT 7
ADH73668
ID   ADH73668 standard; peptide; 11 AA.
XX
AC   ADH73668;
XX
DT   22-APR-2004  (first entry)
XX
DE   Novel recombinant adenovirus-related peptide 2.
XX
KW   adenovirus; fibre protein; target peptide; TP; CD21 receptor; cytostatic;
KW   immunomodulator; antiinflammatory; gene therapy; B lymphocyte; B cell;
KW   leukaemia; lymphoma; immune disorder; inflammation; Epstein Barr Virus.
XX
OS   Human herpesvirus 4.
XX
PN   FR2842823-A1.
XX
PD   30-JAN-2004.
XX
PF   25-JUL-2002; 2002FR-00009426.
XX
PR   25-JUL-2002; 2002FR-00009426.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Dhalluin JC,  Renaut L,  Colin M;
XX
DR   WPI; 2004-135600/14.
XX
PT   Recombinant adenovirus with specific tropism for B cells, useful e.g. for
PT   gene therapy of leukemia, includes a fiber protein that contains a
PT   peptide specific for the CD21 receptor.
XX
PS   Disclosure; Page 6; 24pp; French.
XX
CC   This invention relates to a novel recombinant adenovirus (A) which
CC   contains a sequence encoding a fibre protein. The fibre protein of the
CC   invention contains a target peptide (TP) specific for the CD21 receptor.
CC   The invention may be useful for the development of compounds with a
CC   cytostatic, immunomodulator or antiinflammatory activity or for gene
CC   therapy. The novel adenovirus may be used to transfect genes into B
CC   lymphocytes for experimental, industrial, vaccinating or therapeutic
CC   purposes, particularly for treating diseases associated with B cells,
CC   such as leukaemia, lymphoma, immune disorders and inflammation. The
CC   incorporation of TP provides specific tropism of the virus for B cells.
CC   The present sequence is that of a peptide, responsible for recognition of
CC   CD21 and derived from human herpes virus 4 (Epstein Barr Virus), which is
CC   related to the invention.
XX
SQ   Sequence 11 AA;

  Query Match             65.9%;  Score 60;  DB 8;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 EDPGKQLYNVE 15
              ||||||||||
Db          1 EDPGKQLYNVE 11


RESULT 8
AAR95868
ID   AAR95868 standard; peptide; 11 AA.
XX
AC   AAR95868;
XX
DT   28-OCT-1996  (first entry)
XX
DE   CR2 receptor ligand for intracellular delivery of chemical agents.
XX
KW   CR2; CD21; membrane glycoprotein; B cell; lymphocyte; epithelial;
KW   receptor mediated endocytosis; delivery; targeting; leukaemia; EBV;
KW   Epstein-Barr virus; conjugate.
XX
OS   Synthetic.
XX
PN   WO9608263-A1.
XX
PD   21-MAR-1996.
XX
PF   12-SEP-1995;   95WO-US011515.
XX
PR   13-SEP-1994;   94US-00305770.
XX
```

PA (THER-) THERATECH INC.
XX
PI Ramesh K;
XX
DR WPI; 1996-179718/18.
XX
PT Targeting of chemical agents to CR2(+) cells - using a ligand capable of
PT binding to the CR2 receptor and inducing endocytosis, coupled to a
PT chemical agent, e.g. ricin A.
XX
PS Claim 3; Page 28; 50pp; English.
XX
CC AAR95867-R95871 are ligands of the membrane glycoprotein CR2 receptor
CC (CR2 is also known as CD21) which is found on mature B lymphocytes and
CC certain epithelial cells e.g. cervical epithelium. CR2 is a receptor for
CC Epstein-Barr virus and complement fragments C3d/C3dg. The ligands of this
CC receptor are derived from the N-terminus of the Epstein-Barr virus
CC glycoprotein gp350/220 or the complement component C3dg. The ligands are
CC coupled to a chemical agent for delivery of the agent into a cell bearing
CC the CR2 receptor via receptor-mediated endocytosis. The agent for
CC delivery may be ricin A or other cytotoxic agent for selectively killing
CC CR2 receptor bearing leukaemic B cells or may be agents such as
CC transforming nucleic acids, gene regulators, labels, antigens and drugs
XX
SQ Sequence 11 AA;

Query Match 60.4%; Score 55; DB 2; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.14;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 EDPGKQLYNVE 15
|||||.|||||
Db 1 EDPGKNLYNVE 11

RESULT 9
AAW27141
ID AAW27141 standard; peptide; 11 AA.
XX
AC AAW27141;
XX
DT 23-MAR-1998 (first entry)
XX
DE Complement receptor fragment C3dg ligand.
XX
KW Biodegradable spacer; prodrug; T lymphocyte; endocytosis; cytotoxin;
KW liposome; protease-sensitive; complement receptor C3dg.
XX
OS Homo sapiens.
XX
PN WO9733618-A1.
XX
PD 18-SEP-1997.
XX
PF 12-MAR-1997; 97WO-US003832.
XX
PR 15-MAR-1996; 96US-00616693.
XX
PA (THER-) THERATECH INC.
PA (UTAH) UNIV UTAH RES FOUND.
XX
PI Prakash RK, Kopecek J, Kopeckova P, Omelyanenko V;
XX
DR WPI; 1997-470650/43.
XX
PT Compositions for targetted delivery to T lymphocytes - comprising a water
PT soluble polymer linked via a spacer to a ligand which binds a T cell
PT receptor and to a chemical agent.
XX
PS Disclosure; Page 39; 59pp; English.
XX
CC This sequence represents a complement fragment C3dg derived ligand which
CC was used in a new composition for intracellular delivery of a chemical
CC agent capable of eliciting a selected effect when delivered
CC intracellularly into a T lymphocyte. The composition has the formula [L-
CC S]a-C-[S-A]b where; L = a ligand capable of binding to a receptor on the
CC T lymphocyte and stimulating receptor-mediated endocytosis of the
CC composition; A = the chemical agent; S = a spacer; C = a water soluble
CC polymer having functional groups compatible with forming covalent bonds
CC with the ligand, chemical agent, and spacer; a = an integer of at least 2
CC ; and b = an integer of at least 1. The composition can be used for
CC selectively targeting T lymphocytes with chemical agents such as
CC cytotoxins, transforming nucleic acids, gene regulators, labels, antigens
CC or drugs such as adriamycin. They can be used for treating T-cell-
CC associated diseases such as arthritis, T-cell lymphoma, skin cancers,
CC diseases resulting from HIV infections, or tissue graft rejection
XX
SQ Sequence 11 AA;

Query Match 60.4%; Score 55; DB 2; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.14;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 EDPGKQLYNVE 15
|||||.|||||
Db 1 EDPGKNLYNVE 11

RESULT 10
AAW87720
ID AAW87720 standard; peptide; 11 AA.
XX
AC AAW87720;
XX
DT 09-MAR-1999 (first entry)
XX
DE Epitope involved in CR2 binding.
XX
KW Epitope; viral binding; B lymphocyte EBV receptor; CR2; cell targeting;
KW intra-cellular delivery; T-lymphocyte; cell death.
XX
OS Synthetic.
XX
PN WO9851336-A1.
XX
PD 19-NOV-1998.
XX
PF 04-MAY-1998; 98WO-US009057.
XX
PR 15-MAY-1997; 97US-00857009.
XX
PA (THER-) THERATECH INC.
XX
PI Prakash RK, Kumar V;
XX
DR WPI; 1999-045193/04.
XX
PT Composition for intra-cellular delivery of chemical agent - are capable
PT of eliciting selected effect when delivered into T-lymphocytes.
XX
PS Disclosure; Page 3; 46pp; English.
XX
CC The present sequence represents an epitope that is involved in viral
CC binding to the B lymphocyte EBV receptor (CR2). The peptide acts as a
CC cell targeting moiety, i.e. a ligand, in the composition of the
CC invention. The specification describes a composition for intra-cellular
CC delivery of a chemical agent capable of eliciting a selected effect when
CC delivered into T-lymphocytes. The composition is used to deliver chemical
CC agents in vitro. These agents include a cell targeting moiety, such as
CC growth factor or an antigen binding protein, and they kill cells by
CC mechanisms different from, e.g. conventional chemotherapy
XX
SQ Sequence 11 AA;

Query Match 60.4%; Score 55; DB 2; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.14;

```
  Matches   10;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 EDPGKQLYNVE 15
              ||||| |||||
Db          1 EDPGKNLYNVE 11


RESULT 11
AAW46335
ID   AAW46335 standard; peptide; 10 AA.
XX
AC   AAW46335;
XX
DT   08-MAY-1998  (first entry)
XX
DE   Binding domain of chimeric adenovirus penton base protein.
XX
KW   Integrin; cell surface receptor; penton base protein; adenovirus;
KW   binding site; binding domain; cell surface binding site; gene therapy;
KW   bispecific molecule; antibody; adenoviral transfer vector; pAT.
XX
OS   Synthetic.
XX
PN   US5712136-A.
XX
PD   27-JAN-1998.
XX
PF   17-APR-1996;   96US-00634060.
XX
PR   08-SEP-1994;   94US-00303162.
XX
PA   (GENV-) GENVEC INC.
XX
PI   Bruder JT,  Mcvey DL,  Wickham TJ,  Roelvink PW,  Kovesdi I;
PI   Brough DE;
XX
DR   WPI; 1998-119984/11.
XX
PT   Methods for introducing adenovirus into cells - used for genetic
PT   engineering and gene therapy.
XX
PS   Claim 27; Col 11; 56pp; English.
XX
CC   The present sequence represents a binding domain of a chimeric adenovirus
CC   penton base protein, which is recognised by the CR2 receptor. The penton
CC   base protein binds to cell surface receptors called integrins. The
CC   integrins not only provide a binding site for the adenoviral penton base
CC   protein, but also mediate cellular adhesion to the extracellular matrix
CC   molecules. The specification describes a method of introducing an
CC   adenovirus into a cell in vitro having a particular cell surface binding
CC   site. The adenovirus is contacted with a bispecific molecule (e.g.
CC   bispecific antibody) comprising a component that selectively binds a
CC   binding domain of the penton base protein of the adenovirus and a second
CC   component that selectively binds the cell surface binding site. A complex
CC   of the adenovirus and the bispecific molecule is formed, and the cell is
CC   contacted with it to allow entry of the adenovirus into the cell. The
CC   methods can be used for research and the vectors can be used for gene
CC   therapy
XX
SQ   Sequence 10 AA;

  Query Match             56.0%;  Score 51;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.51;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DPGKQLYNVE 15
              :|||||||||
Db          1 EPGKQLYNVE 10


RESULT 12
AAR95584
ID   AAR95584 standard; peptide; 14 AA.
XX
AC   AAR95584;
XX
DT   27-AUG-2003  (revised)
DT   16-DEC-1996  (first entry)
XX
DE   PepC3 derived from C3d receptor of Epstein Barr virus.
XX
KW   Epstein Barr virus; EBV; gp350; binding agent; CD21; CD11b; CD11c; CD23;
KW   endothelial cell; inhibitor; type II molecule; C-lectin family; antibody;
KW   IgE receptor; haematopoietic cell; histamine; Factor X; therapy; uveitis;
KW   inflammatory disease; autoimmune disease; allergic disease; arthritis;
KW   systemic lupus erythematosus; Mashimotos thyroiditis; multiple sclerosis;
KW   diabetes; dermatitis; inflammatory bowel disease; ulcerative colitis;
KW   Crohn's disease; Sjogren's syndrome; psoriasis; urticaria; insulitis;
KW   nephrotic syndrome; glomerulonephritis; asthma; eczema; bronchitis; COPD;
KW   graft-versus-host disease; chronic lymphocytic leukaemia; rhinitis;
KW   B-cell malignancy; hairy cell leukaemia; pro-inflammatory cytokine;
KW   chronic obstructive pulmonary disease.
XX
OS   Human herpesvirus 4.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   13
FT                   /note= "amidated"
XX
PN   WO9612742-A1.
XX
PD   02-MAY-1996.
XX
PF   20-OCT-1995;   95WO-EP004110.
XX
PR   25-OCT-1994;   94GB-00021463.
PR   20-JUN-1995;   95GB-00012480.
PR   30-JUN-1995;   95GB-00013415.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Bonnefoy JMP,  Lecoanet-Henchoz S;
XX
DR   WPI; 1996-230557/23.
XX
PT   Treatment of inflammatory, auto-immune or allergic diseases - using a
PT   binding agent for CD21, CD11b, CD11c or 70-8 kD or 115 kD proteins
PT   expressed on endothelial cells.
XX
PS   Example 7; Page 25; 52pp; English.
XX
CC   This sequence represents a fragment of the C3d receptor protein
CC   (complement type 2 receptor) of Epstein-barr virus (EBV). The sequence is
CC   a binding agent to CD21, CD11b, CD11c, a 70-85 Kd protein expressed on
CC   endothelial cells, and to a 115 kD protein expressed on endothelial
CC   cells. Binding agents such as this sequence can be used to block the
CC   interaction between CD23 and its binding ligands. CD23 is a type II
CC   molecule of the C-lectin family, and is a low affinity receptor for IgE
CC   expressed on the surface of various haematopoietic cell types. Cellular
CC   activities involving CD23 include regulation of IgE and histamine
CC   release. Other binding agents that can be used include antibodies
CC   (preferably humanised or chimeric), and Factor X, or fragments of these
CC   sequences. The binding agents can be used in the treatment or prophylaxis
CC   of inflammatory, autoimmune, or allergic diseases. These diseases include
CC   arthritis, systemic lupus erythematosus, multiple sclerosis, diabetes,
CC   psoriasis, asthma, chronic obstructive pulmonary disease (COPD), and
CC   bronchitis. The binding agents may also be useful against B-cell
CC   malignancies (such as chronic lymphocytic leukaemia), and for studying
CC   the interactions between CD23 and its ligands. EBV gp350 fragments, and
CC   other binding agents provide effective treatments by suppressing the de
CC   novo synthesis of pro-inflammatory cytokines. (Updated on 27-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 14 AA;

  Query Match             50.5%;  Score 46;  DB 2;  Length 14;
```

```
  Best Local Similarity   100.0%;  Pred. No. 4.3;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 GKQLYNVEA 16
              |||||||||
Db          1 GKQLYNVEA 9


RESULT 13
AAW32826
ID   AAW32826 standard; peptide; 16 AA.
XX
AC   AAW32826;
XX
DT   17-OCT-2003  (revised)
DT   14-JAN-1998  (first entry)
XX
DE   HIV-1 CDC4 envelope glycoprotein 120 T cell epitope T1.
XX
KW   Hydrophilic; antigenic determinant; HIV; envelope; glycoprotein; env; gp;
KW   recognition; B lymphocyte; type specific; antibody; vaccine; protection;
KW   immune response; infection; neutralisation; epitope.
XX
OS   Human immunodeficiency virus 1.
XX
PN   WO9714436-A1.
XX
PD   24-APR-1997.
XX
PF   18-OCT-1996;   96WO-US016911.
XX
PR   20-OCT-1995;   95US-00546515.
PR   09-FEB-1996;   96US-00599266.
XX
PA   (UYDU-) UNIV DUKE.
XX
PI   Haynes BF,  Palker TJ;
XX
DR   WPI; 1997-244862/22.
XX
PT   Synthetic human immunodeficiency virus vaccine - comprising hydrophilic
PT   peptide corresponding to at least 1 antigenic determinant of envelope
PT   glyco:protein recognised by B lymphocytes.
XX
PS   Disclosure; Page 23; 104pp; English.
XX
CC   An essentially pure hydrophilic peptide, comprising at least 1 antigenic
CC   determinant of human immunodeficiency virus (HIV) envelope (env)
CC   glycoprotein (gp) recognised by B lymphocytes, when covalently linked to
CC   a carrier molecule, i.e. the present sequence, induces the production of
CC   high titres of protective, type specific anti-HIV antibodies (Ab) in a
CC   mammal. The peptide can be used in vaccines for producing a protective
CC   immune response to HIV infection, while a HIV neutralising Ab can be
CC   induced in a primate by administering a composition comprising HIV env
CC   peptides that disrupt gp120/gp41 interactions. (Updated on 17-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 16 AA;

  Query Match            41.8%;  Score 38;  DB 2;  Length 16;
  Best Local Similarity   54.5%;  Pred. No. 83;
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              |||:  || :|
Db          5 NRWQVVGKAMY 15


RESULT 14
ABP18544
ID   ABP18544 standard; peptide; 11 AA.
XX
AC   ABP18544;
XX
DT   11-SEP-2003  (revised)
DT   15-JUL-2002  (first entry)
XX
DE   HIV B62 super motif env peptide #119.
XX
KW   HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
KW   vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
KW   vaccine; HIV infection; immunisation; virucide.
XX
OS   Human immunodeficiency virus 1.
XX
PN   WO200124810-A1.
XX
PD   12-APR-2001.
XX
PF   05-OCT-2000; 2000WO-US027766.
XX
PR   05-OCT-1999;   99US-00412863.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S,  Livingston BD,  Chesnut R;
PI   Baker DM,  Celis E,  Kubo RT,  Grey HM;
XX
DR   WPI; 2001-354887/37.
XX
PT   Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT   peptide groups, useful for vaccinating against HIV-1.
XX
PS   Claim 32; Page 249; 448pp; English.
XX
CC   The present invention describes a composition (I) comprising a prepared
CC   human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC   sequence selected from 51 defined amino acid sequences (ABL25347 to
CC   ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
CC   be used for immunising subjects against HIV-1 infections. The use of
CC   group-based vaccines has several advantages over traditional vaccines,
CC   particularly when compared to the use of whole antigens in vaccine
CC   compositions. There is evidence that the immune response to whole
CC   antigens is directed largely toward variable regions of the antigen,
CC   allowing for immune escape due to mutations. The groups for inclusion in
CC   an group-based vaccine may be selected from conserved regions of viral or
CC   tumour-associated antigens, which therefore reduces the likelihood of
CC   escape mutants. Furthermore, immunosuppressive groups that may be present
CC   in whole antigens can be avoided with the use of group-based vaccines. An
CC   additional advantage of an group-based vaccine approach is the ability to
CC   combine selected groups (CTL and HTL), and further, to modify the
CC   composition of the groups, achieving, for example, enhanced
CC   immunogenicity. Accordingly, the immune response can be modulated, as
CC   appropriate, for the target disease. Similar engineering of the response
CC   is not possible with traditional approaches. ABP11501 to ABP25412
CC   represent peptide sequences used in the exemplification of the present
CC   invention. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match            40.7%;  Score 37;  DB 4;  Length 11;
  Best Local Similarity   45.5%;  Pred. No. 82;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          1 NMWQEVGKAMY 11


RESULT 15
AAR24423
ID   AAR24423 standard; peptide; 15 AA.
XX
AC   AAR24423;
XX
DT   25-MAR-2003  (revised)
```

DT 21-NOV-1992 (first entry)
XX
DE Sequence of T helper peptide of gp120 at amino acid residue numbers 421-436(T1).
XX
KW Vaccine; AIDS; HIV-1; carrier peptide.
XX
OS Homo sapiens.
XX
PN WO9208491-A1.
XX
PD 29-MAY-1992.
XX
PF 19-NOV-1991; 91WO-US008653.
XX
PR 20-NOV-1990; 90US-00616247.
XX
PA (TANO-) TANOX BIOSYSTEMS INC.
XX
PI Chang TW, Fung MSC;
XX
DR WPI; 1992-199955/24.
XX
PT Vaccines comprising anti-idiotype antibody conjugates - induce prodn. of
PT neutralising antibodies against HIV-1 for immunisation against HIV
PT infection and AIDS.
XX
PS Claim 12; Page 26 and page 15; 29pp; English.
XX
CC The invention includes epitope-directed immunization with a vaccine in
CC which an anti-idiotype antibody is conjugated to a carrier, which can be
CC either a protein or its derived T helper peptide. The carrier is one
CC against which the vaccine recipient has previously immunized or otherwise
CC previously exposed, or which enhances the immune response against the
CC anti-idiotype antibody. One exemplary anti-idiotype antibody which
CC induces antibodies against the PND is AB19-4. Where the anti-idiotype
CC induces Ab3 against HIV-1, the carrier preferably is HBsAg or HIV-1 p24,
CC or a peptide of either HBsAg or HIV-1 p24 including a T helper
CC determinant. "PND" = the principal neutralizing determinant ("PND") of
CC gp120. A T-helper peptide with the sequence in AAR24423, or immunological
CC equivalents of this sequence is suitable for conjugation with AB19-4 or
CC other anti-idiotypes which induce Ab3 against HIV-1. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
SQ Sequence 15 AA;

Query Match 40.7%; Score 37; DB 2; Length 15;
Best Local Similarity 45.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15
```

Search completed: August 25, 2005, 00:00:08
Job time : 163 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 24, 2005, 23:53:26 ; Search time 38 Seconds
(without alignments)
40.512 Million cell updates/sec

Title: US-09-865-281A-1
Perfect score: 91
Sequence: 1 KNRWEDPGKQLYNVEA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2773

Minimum DB seq length: 0
Maximum DB seq length: 16

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	30.8	14	2	S57569	T cell receptor V-
2	28	30.8	16	2	G24304	ribosomal protein
3	26	28.6	14	2	PH1617	Ig H chain V-D-J r
4	24	26.4	13	1	XAVI9B	angiotensin-conver
5	23	25.3	15	2	PT0091	H+-transporting tw
6	23	25.3	16	2	A31963	pyruvate dehydroge
7	22	24.2	6	2	B35640	cerebellar degener
8	22	24.2	9	2	B20569	serum amyloid P-co
9	22	24.2	10	2	S65388	cytochrome-c oxida
10	22	24.2	10	2	S59625	beta-galactosidase
11	22	24.2	10	2	S77990	cytochrome-c oxida
12	22	24.2	10	2	T17054	cytochrome-c oxida
13	22	24.2	10	2	T14043	cytochrome-c oxida
14	22	24.2	10	2	T14054	cytochrome-c oxida
15	22	24.2	10	2	T17066	cytochrome-c oxida
16	22	24.2	10	2	T17069	cytochrome-c oxida
17	22	24.2	10	2	T12308	cytochrome-c oxida
18	22	24.2	10	2	T12312	cytochrome-c oxida
19	22	24.2	10	2	T12329	cytochrome-c oxida
20	22	24.2	10	2	T12316	cytochrome-c oxida
21	22	24.2	10	2	T12321	cytochrome-c oxida
22	22	24.2	11	2	T17078	cytochrome-c oxida
23	22	24.2	11	2	S07207	Crinia-angiotensin
24	22	24.2	12	2	S10624	lipovitellin - Afr
25	22	24.2	12	2	S21163	NAD ADP-ribosyltra
26	22	24.2	15	2	S61284	phosphoprotein, 80
27	22	24.2	15	2	S43634	cytochrome-c oxida
28	22	24.2	16	2	PT0282	Ig heavy chain CDR
29	21	23.1	9	2	B45796	dihydrolipoamide S
30	21	23.1	13	2	PT0331	Ig heavy chain CRD
31	21	23.1	13	2	S54344	glyceraldehyde-3-p
32	21	23.1	13	2	PC2369	unidentified 85K p
33	21	23.1	14	1	QMWAPP	polistes mastopara
34	21	23.1	15	2	PD0444	coupling factor 6
35	20	22.0	11	2	PT0273	Ig heavy chain CRD
36	20	22.0	13	1	MTCMAD	melanotropin alpha
37	20	22.0	13	1	MTHOAD	melanotropin alpha
38	20	22.0	14	2	PH1614	Ig H chain V-D-J r
39	20	22.0	14	2	PH1623	Ig H chain V-D-J r
40	20	22.0	14	2	PC4376	telomeric and tetr
41	20	22.0	15	2	PT0097	glutathione peroxi
42	20	22.0	15	2	S02381	probable membrane
43	20	22.0	16	1	MTDFBS	melanotropin beta
44	19	20.9	7	2	A44428	platelet aggregati
45	19	20.9	9	1	AKLQIM	locustamyoinhibiti

ALIGNMENTS

RESULT 1
S57569
T cell receptor V-J junctional alpha chain region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C;Accession: S57569
R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.
submitted to the EMBL Data Library, June 1995
A;Description: T cell receptor repertoire for a viral epitope in humans is diversified b
A;Reference number: S57494
A;Accession: S57569
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-14 <BUR>
A;Cross-references: EMBL:Z49955; NID:g887482; PIDN:CAA90226.1; PID:g887483
C;Keywords: T-cell receptor

Query Match 30.8%; Score 28; DB 2; Length 14;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          5 EDPGKQLY 12
              || | | |
Db          5 EDTGNQFY 12
```

RESULT 2
G24304
ribosomal protein H [validated] - Haloarcula marismortui (fragment)
C;Species: Haloarcula marismortui
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 21-Jul-2000
C;Accession: G24304
R;Shoham, M.; Dijk, J.; Reinhardt, R.; Wittmann-Liebold, B.
FEBS Lett. 204, 323-330, 1986
A;Title: Purification and characterization of ribosomal proteins from the 30 S subunit of
A;Reference number: A24304
A;Accession: G24304
A;Molecule type: protein
A;Residues: 1-16 <SHO>
C;Keywords: protein biosynthesis; ribosome

Query Match 30.8%; Score 28; DB 2; Length 16;
Best Local Similarity 55.6%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 PGKQLYNVE 15
              || : || |
Db          1 PGNKYYNDE 9
```

RESULT 3
PH1617

```
Ig H chain V-D-J region (clone B-less 32) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1617
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A;Reference number: PH1580; MUID:93301609; PMID:8315387
A;Accession: PH1617
A;Molecule type: DNA
A;Residues: 1-14 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin

  Query Match             28.6%;  Score 26;  DB 2;  Length 14;
  Best Local Similarity   62.5%;  Pred. No. 5.4e+02;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 KQLYNVEA 16
              :||:|| |
Db          4 RQLFNVYA 11


RESULT 4
XAVI9B
angiotensin-converting enzyme inhibitor V-9 - jararaca
C;Species: Bothrops jararaca (jararaca)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 09-Jul-2004
C;Accession: A01253
R;Ondetti, M.A.; Williams, N.J.; Sabo, E.F.; Pluscec, J.; Weaver, E.R.; Kocy, O.
Biochemistry 10, 4033-4039, 1971
A;Title: Angiotensin-converting enzyme inhibitors from the venom of Bothrops jararaca. I
A;Reference number: A90356; MUID:72118526; PMID:4334402
A;Accession: A01253
A;Molecule type: protein
A;Residues: 1-13 <OND>
A;Cross-references: UNIPROT:P01020
A;Note: the structure of the peptide was confirmed by synthesis
C;Comment: This peptide also potentiates bradykinin by inhibiting the kinases that inact
C;Superfamily: bradykinin-potentiating peptide
C;Keywords: angiotensin-converting enzyme inhibitor; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

  Query Match             26.4%;  Score 24;  DB 1;  Length 13;
  Best Local Similarity   37.5%;  Pred. No. 1.1e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 WEDPGKQL 11
              |  || ::
Db          4 WPRPGPEI 11


RESULT 5
PT0091
H+-transporting two-sector ATPase (EC 3.6.3.14) alpha chain - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 03-Jun-2002
C;Accession: PT0091
R;Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.
submitted to JIPID, July 1998
A;Description: Proteome analysis of mouse brain.
A;Reference number: PT0091
A;Accession: PT0091
A;Molecule type: protein
A;Residues: 1-15 <KAW>
A;Experimental source: brain, striatum
C;Keywords: hydrolase

  Query Match             25.3%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PGKQLY 12
              ||:: |
Db          4 PGREAY 9


RESULT 6
A31963
pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) alpha chain type I - pig roundworm (fragr
C;Species: Ascaris suum (pig roundworm)
C;Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 09-Jul-2004
C;Accession: A31963
R;Thissen, J.; Komuniecki, R.
J. Biol. Chem. 263, 19092-19097, 1988
A;Title: Phosphorylation and inactivation of the pyruvate dehydrogenase from the anaerobi
A;Reference number: A31963; MUID:89066711; PMID:3198613
A;Accession: A31963
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-16 <THI>
A;Cross-references: UNIPROT:P26267
C;Keywords: mitochondrion; oxidoreductase; phosphoprotein

  Query Match             25.3%;  Score 23;  DB 2;  Length 16;
  Best Local Similarity   57.1%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 DPGKQLY 12
              |||   |
Db          9 DPGTSSY 15


RESULT 7
B35640
cerebellar degeneration-related protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 24-Jun-1993
C;Accession: B35640
R;Chen, Y.T.; Rettig, W.J.; Yenamandra, A.K.; Kozak, C.A.; Chaganti, R.S.K.; Posner, J.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 3077-3081, 1990
A;Title: Cerebellar degeneration-related antigen: a highly conserved neuroectodermal mark
A;Reference number: A35640; MUID:90222173; PMID:2326268
A;Accession: B35640
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-6 <CHE>

  Query Match             24.2%;  Score 22;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 WED 6
              |||
Db          2 WED 4


RESULT 8
B20569
serum amyloid P-component - smooth dogfish (fragment)
C;Species: Mustelus canis (smooth dogfish)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
C;Accession: B20569; A05074
R;Robey, F.A.; Tanaka, T.; Liu, T.Y.
J. Biol. Chem. 258, 3889-3894, 1983
A;Title: Isolation and characterization of two major serum proteins from the dogfish, Mus
A;Reference number: A92419; MUID:83160932; PMID:6403520
A;Accession: B20569
A;Molecule type: protein
A;Residues: 1-9 <ROB>
A;Cross-references: UNIPROT:P19095
C;Keywords: amyloid

  Query Match             24.2%;  Score 22;  DB 2;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 2.8e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          7 PGKQL 11
              ||| |
Db          3 PGKSL 7


RESULT 9
S65388
cytochrome-c oxidase (EC 1.9.3.1) chain VII c, hepatic - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S65388; S65389
R;Schaegger, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.
Eur. J. Biochem. 230, 235-241, 1995
A;Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-term
A;Reference number: S65372; MUID:95324529; PMID:7601105
A;Accession: S65388
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <SCH>
A;Cross-references: UNIPROT:P80432
A;Accession: S65389
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <SC2>
C;Superfamily: cytochrome-c oxidase chain VIIc
C;Keywords: oxidoreductase

  Query Match             24.2%;  Score 22;  DB 2;  Length 10;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 EDPGK 9
              | |||
Db          5 EGPGK 9


RESULT 10
S59625
beta-galactosidase alpha chain - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C;Accession: S59625
R;Calugaru, S.V.; Hall, B.G.; Sinnott, M.L.
Biochem. J. 312, 281-286, 1995
A;Title: Catalysis by the large subunit of the second beta-galactosidase of Escherichia
A;Reference number: S59625; MUID:96077156; PMID:7492325
A;Accession: S59625
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <CAL>

  Query Match             24.2%;  Score 22;  DB 2;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 RWED 6
              |||:
Db          3 RWEN 6


RESULT 11
S77990
cytochrome-c oxidase (EC 1.9.3.1) chain VIIc - bigeye tuna (fragment)
C;Species: Thunnus obesus (bigeye tuna)
C;Date: 17-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: S77990
R;Arnold, S.; Lee, J.; Kim, M.; Song, E.; Linder, D.; Lottspeich, F.; Kadenbach, B.
submitted to the Protein Sequence Database, June 1997
A;Reference number: S77980
A;Accession: S77990
A;Molecule type: protein
A;Residues: 1-10 <ARN>
A;Cross-references: UNIPROT:P80982
A;Experimental source: heart; liver
C;Genetics:
A;Genome: nuclear
C;Function:
A;Pathway: oxidative phosphorylation; respiratory chain
C;Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

  Query Match             24.2%;  Score 22;  DB 2;  Length 10;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 EDPGK 9
              | |||
Db          5 EGPGK 9


RESULT 12
T17054
cytochrome-c oxidase (EC 1.9.3.1) chain I - Basiliscus plumifrons mitochondrion (fragmen
C;Species: mitochondrion Basiliscus plumifrons
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17054
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.
J. Mol. Evol. 44, 660-674, 1997
A;Title: Evolutionary shifts in three major structural features of the mitochondrial gen
A;Reference number: Z18674; MUID:97315309; PMID:9169559
A;Accession: T17054
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: UNIPROT:O79888; EMBL:U82680; NID:g3603104; PID:g3603107; PIDN:AAC622
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

  Query Match             24.2%;  Score 22;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NRW 4
              |||
Db          4 NRW 6


RESULT 13
T14043
cytochrome-c oxidase (EC 1.9.3.1) chain I - Lialis jicari mitochondrion (fragment)
C;Species: mitochondrion Lialis jicari
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T14043
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.
Mol. Biol. Evol. 14, 91-104, 1997
A;Title: Two novel gene orders and the role of light-strand replication in rearrangement
A;Reference number: Z17789; MUID:97153826; PMID:9000757
A;Accession: T14043
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: UNIPROT:P92648; EMBL:U71327; NID:g1753244; PID:g1753247; PIDN:AAB482
C;Genetics:
A;Genome: mitochondrion
A;Note: CO1
C;Keywords: mitochondrion; oxidoreductase

  Query Match             24.2%;  Score 22;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NRW 4
              |||
Db          4 NRW 6
```

RESULT 14
T14054
cytochrome-c oxidase (EC 1.9.3.1) chain I - Mabuya aurata mitochondrion (fragment)
C;Species: mitochondrion Mabuya aurata
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T14054
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.
Mol. Biol. Evol. 14, 91-104, 1997
A;Title: Two novel gene orders and the role of light-strand replication in rearrangement
A;Reference number: Z17789; MUID:97153826; PMID:9000757
A;Accession: T14054
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: UNIPROT:P92654; EMBL:U71330; NID:g1753248; PID:g1753251; PIDN:AAB482
C;Genetics:
A;Genome: mitochondrion
A;Note: CO1
C;Keywords: mitochondrion; oxidoreductase

  Query Match 24.2%; Score 22; DB 2; Length 10;
  Best Local Similarity 100.0%; Pred. No. 1.7e+03;
  Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NRW 4
   |||
Db 4 NRW 6

RESULT 15
T17066
cytochrome-c oxidase (EC 1.9.3.1) chain I - Oplurus cuvieri mitochondrion (fragment)
C;Species: mitochondrion Oplurus cuvieri
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17066
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.
J. Mol. Evol. 44, 660-674, 1997
A;Title: Evolutionary shifts in three major structural features of the mitochondrial gen
A;Reference number: Z18674; MUID:97315309; PMID:9169559
A;Accession: T17066
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: UNIPROT:O79903; EMBL:U82685; NID:g3603136; PID:g3603139; PIDN:AAC622
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

  Query Match 24.2%; Score 22; DB 2; Length 10;
  Best Local Similarity 100.0%; Pred. No. 1.7e+03;
  Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NRW 4
   |||
Db 4 NRW 6

Search completed: August 25, 2005, 00:03:48
Job time : 39 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 24, 2005, 23:45:06 ; Search time 171 Seconds
(without alignments)
47.914 Million cell updates/sec

Title: US-09-865-281A-1
Perfect score: 91
Sequence: 1 KNRWEDPGKQLYNVEA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 7514

Minimum DB seq length: 0
Maximum DB seq length: 16

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	25	27.5	8	2	Q94V82	Q94v82 varanus yuw
2	25	27.5	11	2	Q8ADI8	Q8adi8 human immun
3	25	27.5	12	2	Q8QDY4	Q8qdy4 human immun
4	25	27.5	12	2	Q8QDY5	Q8qdy5 human immun
5	25	27.5	12	2	Q8QDY6	Q8qdy6 human immun
6	25	27.5	12	2	Q8QE41	Q8qe41 human immun
7	25	27.5	12	2	Q8QE43	Q8qe43 human immun
8	25	27.5	12	2	Q8QE45	Q8qe45 human immun
9	25	27.5	12	2	Q8QE47	Q8qe47 human immun
10	25	27.5	16	2	Q7DLY3	Q7dly3 solanum tub
11	24	26.4	13	1	BPP1_BOTJA	P01020 bothrops ja
12	24	26.4	14	2	Q6LDN2	Q6ldn2 bacillus st
13	24	26.4	15	2	Q7RBW7	Q7rbw7 plasmodium
14	24	26.4	15	2	Q86865	Q86865 lymphocytic
15	23	25.3	8	2	Q94VA7	Q94va7 varanus sal
16	23	25.3	8	2	Q94VB2	Q94vb2 varanus sal
17	23	25.3	8	2	Q94VB5	Q94vb5 varanus sal
18	23	25.3	9	2	Q94VC6	Q94vc6 varanus pil
19	23	25.3	10	2	Q94VD5	Q94vd5 varanus oli
20	23	25.3	10	2	Q6LBT3	Q6lbt3 mus musculu
21	23	25.3	14	2	Q7S023	Q7s023 neurospora
22	23	25.3	15	2	Q71HK4	Q71hk4 icterus pus
23	23	25.3	15	2	Q71HK5	Q71hk5 icterus gal
24	23	25.3	15	2	Q71HK6	Q71hk6 icterus gal
25	23	25.3	15	2	Q71HL0	Q71hl0 icterus bul
26	22	24.2	9	1	SAMP_MUSCA	P19095 mustelus ca
27	22	24.2	9	2	Q691D6	Q691d6 anolis sagr
28	22	24.2	9	2	Q71DX2	Q71dx2 urostrophus
29	22	24.2	10	1	COXO_RAT	P80432 rattus norv
30	22	24.2	10	1	COXO_THUOB	P80982 thunnus obe
31	22	24.2	10	2	O79885	O79885 anolis pate
32	22	24.2	10	2	O79888	O79888 basiliscus
33	22	24.2	10	2	O79900	O79900 liolaemus p
34	22	24.2	10	2	O79903	O79903 oplurus cuv
35	22	24.2	10	2	O79906	O79906 phrynosoma
36	22	24.2	10	2	P92648	P92648 lialis jica
37	22	24.2	10	2	P92654	P92654 euprepis au
38	22	24.2	10	2	Q8W7U4	Q8w7u4 anolis nite
39	22	24.2	10	2	Q8W8Q2	Q8w8q2 anolis punc
40	22	24.2	10	2	Q8W8Q3	Q8w8q3 anolis nite
41	22	24.2	10	2	Q8W8Q4	Q8w8q4 anolis punc
42	22	24.2	10	2	Q8W916	Q8w916 liolaemus m
43	22	24.2	10	2	Q8W969	Q8w969 anolis orto
44	22	24.2	10	2	Q8W970	Q8w970 anolis nite
45	22	24.2	10	2	Q8W971	Q8w971 anolis fusc

ALIGNMENTS

RESULT 1
Q94V82
ID Q94V82 PRELIMINARY; PRT; 8 AA.
AC Q94V82;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN Name=COI;
OS Varanus yuwonoi.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169856;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:211-226(2001).
DR EMBL; AF407535; AAL10157.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 1045 MW; EFC775A6C3640056 CRC64;

Query Match 27.5%; Score 25; DB 2; Length 8;
Best Local Similarity 60.0%; Pred. No. 1.6e+06;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          3 RWEDP 7
              ||: |
Db          3 RWQSP 7
```

RESULT 2
Q8ADI8
ID Q8ADI8 PRELIMINARY; PRT; 11 AA.
AC Q8ADI8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Truncated vif protein.
GN Name=vif;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22375625; PubMed=12487816; DOI=10.1089/088922202320886325;
RA Harris M.E., Serwadda D., Sewankambo N., Wabwire F., Kim B.,
RA Kigozi G., Kiwanuka N., Phillips J.B., Meehen M., Lutalo T.,
RA Lane J.R., Merling R., Gray R., Wawer M., Birx D.L., Robb M.L.,
RA McCutchan F.E.;
RT "Among 46 near full length HIV type 1 genome sequences from Rakai

```
RT   District, Uganda, subtype D and AD recombinants predominate.";
RL   AIDS Res. Hum. Retroviruses 18:1281-1290(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Harris M.E., Birx D.L., Robb M.L.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Kim B., Phillips J.B., Lane J.R., Merling R., McCutchan F.E.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Lutalo T.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Meehen M., Wawer M.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Serwadda S., Sewankambo N., Wabwire F., Kigozi G., Kiwanuka N.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF484508; AAN73711.1; -.
SQ   SEQUENCE   11 AA;  1492 MW;  75C18E6F82D6C364 CRC64;

  Query Match             27.5%;  Score 25;  DB 2;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWE 5
              :|||:
Db          2 ENRWQ 6


RESULT 3
Q8QDY4
ID   Q8QDY4      PRELIMINARY;      PRT;    12 AA.
AC   Q8QDY4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Vif protein (Fragment).
GN   Name=vif;
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22961413; PubMed=14601597; DOI=10.1089/088922203322493139;
RA   Masharsky A.E., Klimov N.A., Kozlov A.P.;
RT   "Molecular cloning and analysis of full-length genome of HIV type 1
RT   strains prevalent in countries of the former Soviet Union.";
RL   AIDS Res. Hum. Retroviruses 19:933-939(2003).
DR   EMBL; AF413999; AAL78469.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1620 MW;  2A05C18E6F82D6C3 CRC64;

  Query Match             27.5%;  Score 25;  DB 2;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWE 5
              :|||:
Db          2 ENRWQ 6


RESULT 4
Q8QDY5
ID   Q8QDY5      PRELIMINARY;      PRT;    12 AA.
AC   Q8QDY5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Vif protein (Fragment).
GN   Name=vif;
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22961413; PubMed=14601597; DOI=10.1089/088922203322493139;
RA   Masharsky A.E., Klimov N.A., Kozlov A.P.;
RT   "Molecular cloning and analysis of full-length genome of HIV type 1
RT   strains prevalent in countries of the former Soviet Union.";
RL   AIDS Res. Hum. Retroviruses 19:933-939(2003).
DR   EMBL; AF413998; AAL78467.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1620 MW;  2A05C18E6F82D6C3 CRC64;

  Query Match             27.5%;  Score 25;  DB 2;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWE 5
              :|||:
Db          2 ENRWQ 6


RESULT 5
Q8QDY6
ID   Q8QDY6      PRELIMINARY;      PRT;    12 AA.
AC   Q8QDY6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Vif protein (Fragment).
GN   Name=vif;
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22961413; PubMed=14601597; DOI=10.1089/088922203322493139;
RA   Masharsky A.E., Klimov N.A., Kozlov A.P.;
RT   "Molecular cloning and analysis of full-length genome of HIV type 1
RT   strains prevalent in countries of the former Soviet Union.";
RL   AIDS Res. Hum. Retroviruses 19:933-939(2003).
DR   EMBL; AF413997; AAL78465.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1620 MW;  2A05C18E6F82D6C3 CRC64;

  Query Match             27.5%;  Score 25;  DB 2;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWE 5
              :|||:
Db          2 ENRWQ 6


RESULT 6
Q8QE41
ID   Q8QE41      PRELIMINARY;      PRT;    12 AA.
AC   Q8QE41;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Vif protein (Fragment).
GN   Name=vif;
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22961413; PubMed=14601597; DOI=10.1089/088922203322493139;
RA   Masharsky A.E., Klimov N.A., Kozlov A.P.;
```

```
RT   "Molecular cloning and analysis of full-length genome of HIV type 1
RT   strains prevalent in countries of the former Soviet Union.";
RL   AIDS Res. Hum. Retroviruses 19:933-939(2003).
DR   EMBL; AF413972; AAL78402.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1620 MW;  2A05C18E6F82D6C3 CRC64;

  Query Match             27.5%;  Score 25;  DB 2;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWE 5
              :|||:
Db          2 ENRWQ 6


RESULT 7
Q8QE43
ID   Q8QE43      PRELIMINARY;      PRT;    12 AA.
AC   Q8QE43;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Vif protein (Fragment).
GN   Name=vif;
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22961413; PubMed=14601597; DOI=10.1089/088922203322493139;
RA   Masharsky A.E., Klimov N.A., Kozlov A.P.;
RT   "Molecular cloning and analysis of full-length genome of HIV type 1
RT   strains prevalent in countries of the former Soviet Union.";
RL   AIDS Res. Hum. Retroviruses 19:933-939(2003).
DR   EMBL; AF413971; AAL78400.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1620 MW;  2A05C18E6F82D6C3 CRC64;

  Query Match             27.5%;  Score 25;  DB 2;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWE 5
              :|||:
Db          2 ENRWQ 6


RESULT 8
Q8QE45
ID   Q8QE45      PRELIMINARY;      PRT;    12 AA.
AC   Q8QE45;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Vif protein (Fragment).
GN   Name=vif;
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22961413; PubMed=14601597; DOI=10.1089/088922203322493139;
RA   Masharsky A.E., Klimov N.A., Kozlov A.P.;
RT   "Molecular cloning and analysis of full-length genome of HIV type 1
RT   strains prevalent in countries of the former Soviet Union.";
RL   AIDS Res. Hum. Retroviruses 19:933-939(2003).
DR   EMBL; AF413970; AAL78398.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1620 MW;  2A05C18E6F82D6C3 CRC64;

  Query Match             27.5%;  Score 25;  DB 2;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWE 5
              :|||:
Db          2 ENRWQ 6


RESULT 9
Q8QE47
ID   Q8QE47      PRELIMINARY;      PRT;    12 AA.
AC   Q8QE47;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Vif protein (Fragment).
GN   Name=vif;
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22961413; PubMed=14601597; DOI=10.1089/088922203322493139;
RA   Masharsky A.E., Klimov N.A., Kozlov A.P.;
RT   "Molecular cloning and analysis of full-length genome of HIV type 1
RT   strains prevalent in countries of the former Soviet Union.";
RL   AIDS Res. Hum. Retroviruses 19:933-939(2003).
DR   EMBL; AF413969; AAL78396.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1648 MW;  28D5C18E6F82D6C3 CRC64;

  Query Match             27.5%;  Score 25;  DB 2;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KNRWE 5
              :|||:
Db          2 ENRWQ 6


RESULT 10
Q7DLY3
ID   Q7DLY3      PRELIMINARY;      PRT;    16 AA.
AC   Q7DLY3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Beta-fructofuranosidase (Invertase) (EC 3.2.1.26) (Fragment).
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4113;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96279736; PubMed=8710506; DOI=10.1093/nar/24.12.2347;
RA   Bournay A.S., Hedley P.E., Maddison A., Waugh R., Machray G.C.;
RT   "Exon skipping induced by cold stress in a potato invertase gene
RT   transcript.";
RL   Nucleic Acids Res. 24:2347-2351(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Maddison A.L.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X95820; CAA65086.1; -.
DR   GO; GO:0004564; F:beta-fructofuranosidase activity; IEA.
DR   GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
KW   Glycosidase; Hydrolase.
FT   NON_TER       1      1
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1894 MW;  003053E73810C336 CRC64;

  Query Match             27.5%;  Score 25;  DB 2;  Length 16;
```

```
  Best Local Similarity   41.7%;  Pred. No. 4.2e+03;
  Matches    5;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 KNRWEDPGKQLY 12
              ||   ||   :|
Db          2 KNWINDPNAPMY 13


RESULT 11
BPP1_BOTJA
ID   BPP1_BOTJA     STANDARD;      PRT;    13 AA.
AC   P01020; P30421;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Bradykinin-potentiating peptide S3,1 (13A) (Angiotensin-converting
DE   enzyme inhibitor V-9).
OS   Bothrops jararaca (Jararaca), and
OS   Bothrops insularis (Island jararaca) (Queimada jararaca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Bothrops.
OX   NCBI_TaxID=8724, 8723;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=B.jararaca; TISSUE=Venom;
RX   MEDLINE=72118526; PubMed=4334402;
RA   Ondetti M.A., Williams N.J., Sabo E.F., Pluscec J., Weaver E.R.,
RA   Kocy O.;
RT   "Angiotensin-converting enzyme inhibitors from the venom of Bothrops
RT   jararaca. Isolation, elucidation of structure, and synthesis.";
RL   Biochemistry 10:4033-4039(1971).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=B.insularis; TISSUE=Venom;
RX   MEDLINE=90351557; PubMed=2386615;
RA   Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT   "Primary structure and biological activity of bradykinin potentiating
RT   peptides from Bothrops insularis snake venom.";
RL   J. Protein Chem. 9:221-227(1990).
CC   -!- FUNCTION: This peptide both inhibits the activity of the
CC       angiotensin-converting enzyme and enhances the action of
CC       bradykinin by inhibiting the kinases that inactivate it. It acts
CC       as an indirect hypotensive agent.
DR   PIR; A01253; XAVI9B.
KW   Direct protein sequencing; Hypotensive agent;
KW   Pyrrolidone carboxylic acid.
FT   MOD_RES       1      1       Pyrrolidone carboxylic acid.
SQ   SEQUENCE   13 AA;  1388 MW;  6824FC97D83D6774 CRC64;

  Query Match             26.4%;  Score 24;  DB 1;  Length 13;
  Best Local Similarity   37.5%;  Pred. No. 4.9e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 WEDPGKQL 11
              |  || ::
Db          4 WPRPGPEI 11


RESULT 12
Q6LDN2
ID   Q6LDN2      PRELIMINARY;      PRT;    14 AA.
AC   Q6LDN2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   B.stearothermophilus (strain 799) alpha-amylase (B.stearothermophilus
DE   (strain DY-5) alpha-amylase) (Fragment).
OS   Bacillus stearothermophilus.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX   NCBI_TaxID=1422;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88139156; PubMed=3257753;
RA   Satoh H., Nishida H., Isono K.;
RT   "Evidence for movement of the alpha-amylase gene into two
RT   phylogenetically distant Bacillus stearothermophilus strains.";
RL   J. Bacteriol. 170:1034-1040(1988).
DR   EMBL; M29578; AAA22228.1; -.
DR   EMBL; M29577; AAA22225.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1786 MW;  7634F11A1FBF066B CRC64;

  Query Match             26.4%;  Score 24;  DB 2;  Length 14;
  Best Local Similarity   60.0%;  Pred. No. 5.4e+03;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 RWEDP 7
              || :|
Db          4 RWTEP 8


RESULT 13
Q7RBW7
ID   Q7RBW7      PRELIMINARY;      PRT;    15 AA.
AC   Q7RBW7;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   Name=PY06019;
OS   Plasmodium yoelii yoelii.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=73239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=17XNL;
RX   PubMed=12368865; DOI=10.1038/nature01099;
RA   Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA   Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA   Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA   Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA   Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA   Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA   Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA   van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA   Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA   Carucci D.J.;
RT   "Genome sequence and comparative analysis of the model rodent malaria
RT   parasite Plasmodium yoelii yoelii.";
RL   Nature 419:512-519(2002).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABL01001991; EAA18155.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1856 MW;  8512319D8F8CDA96 CRC64;

  Query Match             26.4%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 5.8e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 KQLYN 13
              ||:||
Db          1 KQVYN 5


RESULT 14
Q86865
ID   Q86865      PRELIMINARY;      PRT;    15 AA.
AC   Q86865;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LCMV viral protein protein (Fragment).
```

```
GN   Name=LCMV viral protein;
OS   Lymphocytic choriomeningitis virus.
OC   Viruses; ssRNA negative-strand viruses; Arenaviridae; Arenavirus;
OC   Old world arenaviruses.
OX   NCBI_TaxID=11623;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95190990; PubMed=7533851;
RA   Moskophidis D., Zinkernagel R.M.;
RT   "Immunobiology of cytotoxic T-cell escape mutants of lymphocytic
RT   choriomeningitis virus.";
RL   J. Virol. 69:2187-2193(1995).
DR   EMBL; S75741; AAB33667.1; -.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1599 MW;  2D3720F4F776C1A7 CRC64;

  Query Match             26.4%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5.8e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 EDPG 8
              ||||
Db          5 EDPG 8


RESULT 15
Q94VA7
ID   Q94VA7      PRELIMINARY;      PRT;     8 AA.
AC   Q94VA7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cytochrome c oxidase subunit I (Fragment).
GN   Name=COI;
OS   Varanus salvator salvator.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX   NCBI_TaxID=169831;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ast J.C.;
RT   "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL   Cladistics 17:211-226(2001).
DR   EMBL; AF407526; AAL10130.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  992 MW;  EFC775A5A36411A6 CRC64;

  Query Match             25.3%;  Score 23;  DB 2;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+06;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 RWEDP 7
              ||  |
Db          3 RWSSP 7


Search completed: August 25, 2005, 00:03:04
Job time : 173 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 24, 2005, 23:54:41 ; Search time 40 Seconds
(without alignments)
29.860 Million cell updates/sec

Title: US-09-865-281A-1
Perfect score: 91
Sequence: 1 KNRWEDPGKQLYNVEA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 171351

Minimum DB seq length: 0
Maximum DB seq length: 16

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	100.0	16	3	US-09-070-907-1	Sequence 1, Appli
2	60	65.9	11	4	US-09-039-060A-6	Sequence 6, Appli
3	60	65.9	11	5	PCT-US94-01234-37	Sequence 37, Appl
4	60	65.9	11	5	PCT-US94-01263-7	Sequence 7, Appli
5	51	56.0	10	1	US-08-634-060-33	Sequence 33, Appl
6	51	56.0	10	2	US-08-700-846-5	Sequence 5, Appli
7	37	40.7	15	5	PCT-US92-06688-5	Sequence 5, Appli
8	37	40.7	15	5	PCT-US92-06688-21	Sequence 21, Appl
9	37	40.7	16	1	US-08-213-124-5	Sequence 5, Appli
10	37	40.7	16	1	US-08-488-252-35	Sequence 35, Appl
11	37	40.7	16	2	US-07-847-311A-15	Sequence 15, Appl
12	37	40.7	16	3	US-09-046-373-1	Sequence 1, Appli
13	37	40.7	16	4	US-09-009-953-230	Sequence 230, App
14	37	40.7	16	4	US-09-340-798A-40	Sequence 40, Appl
15	37	40.7	16	4	US-09-311-784A-308	Sequence 308, App
16	37	40.7	16	4	US-09-724-961-51	Sequence 51, Appl
17	37	40.7	16	4	US-09-580-018-51	Sequence 51, Appl
18	37	40.7	16	4	US-09-724-551-51	Sequence 51, Appl
19	36	39.6	13	6	5310729-8	Patent No. 5310729
20	36	39.6	13	6	5310729-8	Patent No. 5310729
21	35	38.5	7	6	5310729-12	Patent No. 5310729
22	35	38.5	7	6	5310729-12	Patent No. 5310729
23	35	38.5	9	6	5310729-39	Patent No. 5310729
24	35	38.5	9	6	5310729-40	Patent No. 5310729
25	35	38.5	9	6	5310729-39	Patent No. 5310729
26	35	38.5	9	6	5310729-40	Patent No. 5310729
27	34	37.4	15	1	US-08-709-047-9	Sequence 9, Appli
28	34	37.4	15	1	US-08-410-360-9	Sequence 9, Appli
29	34	37.4	15	1	US-08-707-801A-9	Sequence 9, Appli
30	34	37.4	15	1	US-08-709-006-9	Sequence 9, Appli
31	34	37.4	15	1	US-08-711-175-9	Sequence 9, Appli
32	34	37.4	15	2	US-08-937-102-26	Sequence 26, Appl
33	34	37.4	15	2	US-08-937-102-27	Sequence 27, Appl
34	34	37.4	15	2	US-08-937-102-28	Sequence 28, Appl
35	34	37.4	15	3	US-08-089-990-1	Sequence 1, Appli
36	33	36.3	13	1	US-08-548-540-153	Sequence 153, App
37	33	36.3	13	5	PCT-US96-09809-153	Sequence 153, App
38	32	35.2	12	2	US-08-448-603A-25	Sequence 25, Appl
39	32	35.2	12	3	US-09-134-075-25	Sequence 25, Appl
40	32	35.2	12	3	US-09-492-739-25	Sequence 25, Appl
41	32	35.2	12	4	US-09-966-931A-25	Sequence 25, Appl
42	32	35.2	15	3	US-09-184-938-7	Sequence 7, Appli
43	31	34.1	9	1	US-08-634-060-32	Sequence 32, Appl
44	31	34.1	9	1	US-08-366-522A-1	Sequence 1, Appli
45	31	34.1	9	2	US-08-700-846-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-070-907-1
; Sequence 1, Application US/09070907
; Patent No. 6238667
; GENERAL INFORMATION:
; APPLICANT: Kohler, Heinz
; TITLE OF INVENTION: METHOD OF AFFINITY CROSS-LINKING BIOLOGICALLY ACTIVE
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES TO ANTIBODIES.
; FILE REFERENCE: 35629
; CURRENT APPLICATION NUMBER: US/09/070,907
; CURRENT FILING DATE: 1998-05-04
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: AMINO ACID
; OTHER INFORMATION: SEQUENCE DERIVED FROM Cd3 peptide
US-09-070-907-1

Query Match 100.0%; Score 91; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 1 KNRWEDPGKQLYNVEA 16

RESULT 2
US-09-039-060A-6
; Sequence 6, Application US/09039060A
; Patent No. 6613563
; GENERAL INFORMATION:
; APPLICANT: Sosnowski, Barbara A.
; APPLICANT: Baird, Andrew
; APPLICANT: Pierce, Glenn F.
; APPLICANT: Curiel, David T.
; APPLICANT: Douglas, Joanne T.
; APPLICANT: Rogers, Buck E.
; TITLE OF INVENTION: VIRAL VECTORS WITH MODIFIED TROPISM
; FILE REFERENCE: 760100.427
; CURRENT APPLICATION NUMBER: US/09/039,060A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 11

```
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Polypeptide capable of targeting receptors such as
;   OTHER INFORMATION: the CR2 receptor
US-09-039-060A-6

  Query Match             65.9%;  Score 60;  DB 4;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.00078;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 EDPGKQLYNVE 15
              ||||||||||
Db          1 EDPGKQLYNVE 11


RESULT 3
PCT-US94-01234-37
; Sequence 37, Application PC/TUS9401234
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  METHODS FOR PRODUCING POLYPEPTIDE
;    TITLE OF INVENTION:  BINDING SITES
;    NUMBER OF SEQUENCES:  76
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/01234
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/084,542
;      FILING DATE:  28-JUN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/012,566
;      FILING DATE:  02-FEB-1993
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
PCT-US94-01234-37

  Query Match             65.9%;  Score 60;  DB 5;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.00078;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 EDPGKQLYNVE 15
              ||||||||||
Db          1 EDPGKQLYNVE 11


RESULT 4
PCT-US94-01263-7
; Sequence 7, Application PC/TUS9401263
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  TARGETING AND DELIVERY OF GENES AND
;    TITLE OF INVENTION:  ANTIVIRAL AGENTS INTO CELLS BY THE ADENOVIRUS PENTON
;    NUMBER OF SEQUENCES:  7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/01263
;      FILING DATE:  03-FEB-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/015,225
;      FILING DATE:  09-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/046,159
;      FILING DATE:  13-APR-1993
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
PCT-US94-01263-7

  Query Match             65.9%;  Score 60;  DB 5;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.00078;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 EDPGKQLYNVE 15
              ||||||||||
Db          1 EDPGKQLYNVE 11


RESULT 5
US-08-634-060-33
; Sequence 33, Application US/08634060
; Patent No. 5712136
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Kovesdi, Imre
;    APPLICANT:  Roelvink, Petrus W.
;    TITLE OF INVENTION:  ADENOVIRAL-MEDIATED CELL TARGETING COMMANDED BY
;    TITLE OF INVENTION:  THE ADENOVIRUS PENTON BASE PROTEIN
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,060
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/303,162
;      FILING DATE:  08-SEP-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kilyk, John Jr.
;      REGISTRATION NUMBER:  30763
;      REFERENCE/DOCKET NUMBER:  71602
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5600
;      TELEFAX:  (312) 616-5700
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-060-33

  Query Match             56.0%;  Score 51;  DB 1;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          6 DPGKQLYNVE 15
              :|||||||||
Db          1 EPGKQLYNVE 10


RESULT 6
US-08-700-846-5
; Sequence 5, Application US/08700846
; Patent No. 5962311
;  GENERAL INFORMATION:
;    APPLICANT:  WICKHAM, THOMAS J.
;    APPLICANT:  ROELVINK, PETRUS W.
;    APPLICANT:  KOVESDI, IMRE
;    TITLE OF INVENTION:  A SHORT-SHAFTED ADENOVIRAL FIBER AND ITS
;    TITLE OF INVENTION:  USE
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LEYDIG, VOIT & MAYER, LTD.
;      STREET:  TWO PRUDENTIAL PLAZA, SUITE 4900
;      CITY:  CHICAGO
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60601-6780
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,846
;      FILING DATE:  21-AUG-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LARCHER, CAROL
;      REGISTRATION NUMBER:  35243
;      REFERENCE/DOCKET NUMBER:  74294
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5600
;      TELEFAX:  (312) 616-5700
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-700-846-5

  Query Match             56.0%;  Score 51;  DB 2;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 DPGKQLYNVE 15
              :|||||||||
Db          1 EPGKQLYNVE 10


RESULT 7
PCT-US92-06688-5
; Sequence 5, Application PC/TUS9206688
;  GENERAL INFORMATION:
;    APPLICANT:  REPLIGEN CORPORATION
;    APPLICANT:  THE ROCKEFELLER UNIVERSITY
;    TITLE OF INVENTION:  MULTIPLE ANTIGEN PEPTIDES FOR USE AS HIV
;    TITLE OF INVENTION:  VACCINES
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  IBM P.C. DOS (Version 3.30)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06688
;      FILING DATE:  19920811
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  744,281
;      FILING DATE:  13 August 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00231/052WO1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-06688-5

  Query Match             40.7%;  Score 37;  DB 5;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 8.9;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          4 NMWQEVGKAMY 14


RESULT 8
PCT-US92-06688-21
; Sequence 21, Application PC/TUS9206688
;  GENERAL INFORMATION:
;    APPLICANT:  REPLIGEN CORPORATION
;    APPLICANT:  THE ROCKEFELLER UNIVERSITY
;    TITLE OF INVENTION:  MULTIPLE ANTIGEN PEPTIDES FOR USE AS HIV
;    TITLE OF INVENTION:  VACCINES
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  IBM P.C. DOS (Version 3.30)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06688
;      FILING DATE:  19920811
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  744,281
;      FILING DATE:  13 August 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00231/052WO1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
```

```
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-06688-21

  Query Match             40.7%;  Score 37;  DB 5;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 8.9;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          4 NMWQEVGKAMY 14


RESULT 9
US-08-213-124-5
; Sequence 5, Application US/08213124
; Patent No. 5693325
;  GENERAL INFORMATION:
;    APPLICANT:  Kahn, Michael
;    TITLE OF INVENTION:  PEPTIDE VACCINES AND METHODS RELATING
;    TITLE OF INVENTION:  THERETO
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/213,124
;      FILING DATE:  15-MAR-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hermanns, Karl R.
;      REGISTRATION NUMBER:  33,507
;      REFERENCE/DOCKET NUMBER:  670063.411
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;      TELEX:  3723836 SEEDANDBERRY
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-213-124-5

  Query Match             40.7%;  Score 37;  DB 1;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 9.5;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 10
US-08-488-252-35
; Sequence 35, Application US/08488252
; Patent No. 5763160
;  GENERAL INFORMATION:
;    APPLICANT:  Chang Yi Wang
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDES AND PROCESS
;    TITLE OF INVENTION:  OF USING SAME FOR THE DETECTION OF ANTIBODIES TO
;    TITLE OF INVENTION:  HUMAN IMMUNODEFICIENCY VIRUS (HIV) GP120 ENVELOPE
;    TITLE OF INVENTION:  PROTEIN, DIAGNOSIS OF AIDS AND PRE-AIDS CONDITIONS
;    TITLE OF INVENTION:  AND AS VACCINES
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVE.
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,252
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08\326,676
;      FILING DATE:  07-Jun-1995
;      APPLICATION NUMBER:  07\726,605
;      FILING DATE:  09-July-1991
;      APPLICATION NUMBER:  07\663,262
;      FILING DATE:  01-Mar-1991
;      APPLICATION NUMBER:  07\155,321
;      FILING DATE:  12-Feb-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C. H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4004 US4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  Amino acids
;      STRANDEDNESS:
;      TOPOLOGY:  Unknown
US-08-488-252-35

  Query Match             40.7%;  Score 37;  DB 1;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 9.5;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 11
US-07-847-311A-15
; Sequence 15, Application US/07847311A
; Patent No. 5976541
;  GENERAL INFORMATION:
;    APPLICANT:  Berzofsky, Jay A.
;    APPLICANT:  Takeshita, Toshiyuki
;    APPLICANT:  Shirai, Mutsunori
;    APPLICANT:  Pendleton, C.D.
;    APPLICANT:  Koslowski, Steven
;    APPLICANT:  Margulies, David H.
;    TITLE OF INVENTION:  Potent Peptide for Stimulation of
;    TITLE OF INVENTION:  Cytotoxic T Lymphocytes Specific for the HIV-I Envelope
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolash & Birch
;      STREET:  301 N. Washington
;      CITY:  Falls Church
```

```
;       STATE:  Virginia
;       COUNTRY:  USA
;       ZIP:  22046-0747
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/847,311A
;       FILING DATE:  06-MAR-1992
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Svensson, Leonard R.
;       REGISTRATION NUMBER:  30,330
;       REFERENCE/DOCKET NUMBER:  1173-392P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-241-1300
;       TELEFAX:  703-241-2848
;   INFORMATION FOR SEQ ID NO:  15:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  16 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  NO
;     FRAGMENT TYPE:  internal
;     ORIGINAL SOURCE:
;       ORGANISM:  Human Immunodeficiency Virus Type I
;     FEATURE:
;       NAME/KEY:  Peptide
;       LOCATION:  1..16
;       OTHER INFORMATION:  /label= peptide
;       OTHER INFORMATION:  /note= "peptide T1, T-cell helper determinant in
;       OTHER INFORMATION:  HIV-I envelope glycoprote..."
US-07-847-311A-15

  Query Match             40.7%;  Score 37;  DB 2;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 9.5;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 12
US-09-046-373-1
; Sequence 1, Application US/09046373
; Patent No. 6235714
; GENERAL INFORMATION:
;  APPLICANT: Sudhir Paul
;  APPLICANT: Larry J. Smith
;  APPLICANT: Gennady Gololobov
;  TITLE OF INVENTION: Methods for Identifying Inducers and
;  TITLE OF INVENTION: Inhibitors of Catalytic Antibodies, Compositions and Their
;  TITLE OF INVENTION: Use
;  FILE REFERENCE: UNMC 63123
;  CURRENT APPLICATION NUMBER: US/09/046,373
;  CURRENT FILING DATE: 1998-03-23
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Human Immunodeficiency Virus-1
US-09-046-373-1

  Query Match             40.7%;  Score 37;  DB 3;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 9.5;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 13
US-09-009-953-230
; Sequence 230, Application US/09009953
; Patent No. 6413517
;   GENERAL INFORMATION:
;       APPLICANT: Sette, Alessandro
;       TITLE OF INVENTION: Identification of Broadly
;                           Reactive DR Restricted Epitopes
;       NUMBER OF SEQUENCES: 274
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Townsend and Townsend and Crew LLP
;            STREET: Two Embarcadero Center, Eighth Floor
;            CITY: San Francisco
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94111-3834
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/009,953
;            FILING DATE: 21-Jan-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 60/036,713
;            FILING DATE: 23-JAN-1997
;            APPLICATION NUMBER: US 60/037,432
;            FILING DATE: 07-FEB-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Weber, Ellen Lauver
;            REGISTRATION NUMBER: 32,762
;            REFERENCE/DOCKET NUMBER: 018623-011520US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 415-576-0200
;            TELEFAX: 415-576-0300
;            TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 230:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 16 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 230:
US-09-009-953-230

  Query Match             40.7%;  Score 37;  DB 4;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 9.5;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          6 NMWQEVGKAMY 16


RESULT 14
US-09-340-798A-40
; Sequence 40, Application US/09340798A
; Patent No. 6534312
;   GENERAL INFORMATION:
;       APPLICANT: SHIVER, JOHN W.
;                  LIU, MARGARET A.
;                  PERRY, HELEN C.
;                  DAVIES, MARY-ELLEN M.
;                  FREED, DANIEL C.
;       TITLE OF INVENTION: VACCINES COMPRISING SYNTHETIC GENES
;       NUMBER OF SEQUENCES: 53
```

```
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
;             STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
;             CITY: RAHWAY
;             STATE: NEW JERSEY
;             COUNTRY: US
;             ZIP: 07065-0907
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/340,798A
;             FILING DATE: 28-Jun-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/877,418
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: HAND, J. MARK
;             REGISTRATION NUMBER: 36,545
;             REFERENCE/DOCKET NUMBER: 19729Y
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 908-594-3905
;             TELEFAX: 908-594-4720
;    INFORMATION FOR SEQ ID NO: 40:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 16 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-340-798A-40

  Query Match             40.7%;  Score 37;  DB 4;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 9.5;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 15
US-09-311-784A-308
; Sequence 308, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
;  APPLICANT: Fikes, John D.
;  APPLICANT:  Hermanson, Gary G.
;  APPLICANT:  Sette, Alessandro
;  APPLICANT:  Ishioka, Glenn Y.
;  APPLICANT:  Livingston, Brian
;  APPLICANT:  Chesnut, Robert W.
;  APPLICANT:  Epimmune Inc.
;  TITLE OF INVENTION: Expression Vectors for Stimulating an
;  TITLE OF INVENTION:  Immune Response and Methods of Using the Same
;  FILE REFERENCE: 39963-20022.01
;  CURRENT APPLICATION NUMBER: US/09/311,784A
;  CURRENT FILING DATE:  1999-05-13
;  PRIOR APPLICATION NUMBER: US 60/085,751
;  PRIOR FILING DATE: 1998-05-15
;  NUMBER OF SEQ ID NOS: 463
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 308
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: HIV1 ENV 566 (peptide F091.15)
US-09-311-784A-308

  Query Match             40.7%;  Score 37;  DB 4;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 9.5;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          6 NMWQEVGKAMY 16


Search completed: August 25, 2005, 00:04:34
Job time : 40 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 25, 2005, 00:03:13 ; Search time 157 Seconds
(without alignments)
39.907 Million cell updates/sec

Title: US-09-865-281A-1
Perfect score: 91
Sequence: 1 KNRWEDPGKQLYNVEA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 307307

Minimum DB seq length: 0
Maximum DB seq length: 16

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	100.0	16	10	US-09-865-281A-1	Sequence 1, Appli
2	91	100.0	16	17	US-10-795-081A-1	Sequence 1, Appli
3	60	65.9	11	15	US-10-408-849-6	Sequence 6, Appli
4	37	40.7	15	14	US-10-133-210-251	Sequence 251, App
5	37	40.7	16	9	US-09-775-805-44	Sequence 44, Appl
6	37	40.7	16	9	US-09-775-805-67	Sequence 67, Appl
7	37	40.7	16	9	US-09-775-805-75	Sequence 75, Appl
8	37	40.7	16	9	US-09-775-805-89	Sequence 89, Appl
9	37	40.7	16	9	US-09-862-849-1	Sequence 1, Appli
10	37	40.7	16	9	US-09-894-018-199	Sequence 199, App
11	37	40.7	16	10	US-09-894-594-66	Sequence 66, Appl
12	37	40.7	16	13	US-10-103-395-230	Sequence 230, App
13	37	40.7	16	14	US-10-114-716A-1	Sequence 1, Appli
14	37	40.7	16	14	US-10-041-414-42	Sequence 42, Appl
15	37	40.7	16	15	US-10-371-525-308	Sequence 308, App
16	37	40.7	16	15	US-10-371-069-308	Sequence 308, App
17	37	40.7	16	15	US-10-371-645-308	Sequence 308, App
18	37	40.7	16	15	US-10-371-260-308	Sequence 308, App
19	37	40.7	16	15	US-10-369-121-40	Sequence 40, Appl
20	37	40.7	16	15	US-10-372-111-9	Sequence 9, Appli
21	37	40.7	16	16	US-10-699-517-11	Sequence 11, Appl
22	37	40.7	16	16	US-10-753-339-44	Sequence 44, Appl
23	37	40.7	16	16	US-10-753-339-67	Sequence 67, Appl
24	37	40.7	16	16	US-10-753-339-75	Sequence 75, Appl
25	37	40.7	16	16	US-10-753-339-89	Sequence 89, Appl
26	37	40.7	16	16	US-10-771-174A-10	Sequence 10, Appl
27	37	40.7	16	16	US-10-889-999-51	Sequence 51, Appl
28	37	40.7	16	16	US-10-890-070-51	Sequence 51, Appl
29	37	40.7	16	16	US-10-474-960A-199	Sequence 199, App
30	37	40.7	16	16	US-10-890-000-51	Sequence 51, Appl
31	37	40.7	16	17	US-10-823-463-51	Sequence 51, Appl
32	37	40.7	16	17	US-10-915-214-11	Sequence 11, Appl
33	37	40.7	16	17	US-10-822-968-51	Sequence 51, Appl
34	37	40.7	16	17	US-10-777-792-51	Sequence 51, Appl
35	37	40.7	16	18	US-10-890-071-51	Sequence 51, Appl
36	37	40.7	16	18	US-10-930-548-1	Sequence 1, Appli
37	37	40.7	16	18	US-10-698-099-11	Sequence 11, Appl
38	37	40.7	16	18	US-10-890-024-51	Sequence 51, Appl
39	37	40.7	16	20	US-11-058-757-51	Sequence 51, Appl
40	34	37.4	16	9	US-09-911-838-225	Sequence 225, App
41	32	35.2	8	17	US-10-937-912-53	Sequence 53, Appl
42	32	35.2	9	9	US-09-854-122-29	Sequence 29, Appl
43	32	35.2	10	17	US-10-937-912-65	Sequence 65, Appl
44	32	35.2	12	10	US-09-966-931-25	Sequence 25, Appl
45	32	35.2	12	16	US-10-459-121-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-09-865-281A-1
; Sequence 1, Application US/09865281A
; Publication No. US20030103984A1
; GENERAL INFORMATION:
; APPLICANT: Kohler, Heinz
; TITLE OF INVENTION: FUSION PROTEINS OF BIOLOGICALLY ACTIVE PEPTIDES AND ANTIBODIES
; FILE REFERENCE: 411.35629PC2
; CURRENT APPLICATION NUMBER: US/09/865,281A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 09/070,907
; PRIOR FILING DATE: 1998-05-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Synthesized peptide with sequence derived from position 1217-1232
US-09-865-281A-1

Query Match 100.0%; Score 91; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 1 KNRWEDPGKQLYNVEA 16

RESULT 2

US-10-795-081A-1
; Sequence 1, Application US/10795081A
; Publication No. US20050033033A1
; GENERAL INFORMATION:
; APPLICANT: Kohler, Heinz
; TITLE OF INVENTION: TRANS-MEMBRANE-ANTIBODY INDUCED INHIBITION OF APOPTOSIS
; FILE REFERENCE: 411.35629AP3
; CURRENT APPLICATION NUMBER: US/10/795,081A
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: 60/451,980
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 09/865,281
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 09/070,907
; PRIOR FILING DATE: 1998-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Synthesized peptide with sequence derived from position 1217-1232
US-10-795-081A-1

Query Match 100.0%; Score 91; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNRWEDPGKQLYNVEA 16
||||||||||||||||
Db 1 KNRWEDPGKQLYNVEA 16

RESULT 3
US-10-408-849-6
; Sequence 6, Application US/10408849
; Publication No. US20040029280A1
; GENERAL INFORMATION:
; APPLICANT: Sosnowski, Barbara A.
; APPLICANT: Baird, Andrew
; APPLICANT: Pierce, Glenn F.
; APPLICANT: Curiel, David T.
; APPLICANT: Douglas, Joanne T.
; APPLICANT: Rogers, Buck E.
; TITLE OF INVENTION: VIRAL VECTORS WITH MODIFIED TROPISM
; FILE REFERENCE: 760100.427C1
; CURRENT APPLICATION NUMBER: US/10/408,849
; CURRENT FILING DATE: 2003-04-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polypeptide capable of targeting receptors such as
; OTHER INFORMATION: the CR2 receptor
US-10-408-849-6

Query Match 65.9%; Score 60; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 EDPGKQLYNVE 15
|||||||||||
Db 1 EDPGKQLYNVE 11

RESULT 4
US-10-133-210-251
; Sequence 251, Application US/10133210
; Publication No. US20030103964A1
; GENERAL INFORMATION:
; APPLICANT: DeLisi, Charles
; APPLICANT: Berzofsky, Jay
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Vaccaro, Dennis
; APPLICANT: Weng, Zhiping
; APPLICANT: Zhang, Chao
; TITLE OF INVENTION: METHODS FOR DESIGNING MOLECULAR CONJUGATES AND
; TITLE OF INVENTION: COMPOSITIONS THEREOF
; FILE REFERENCE: BU-035AX
; CURRENT APPLICATION NUMBER: US/10/133,210
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 281
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 251
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-133-210-251

Query Match 40.7%; Score 37; DB 14; Length 15;
Best Local Similarity 45.5%; Pred. No. 67;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 NRWEDPGKQLY 12
| |:: || :|
Db 5 NMWQEVGKAMY 15

RESULT 5
US-09-775-805-44
; Sequence 44, Application US/09775805
; Publication No. US20010036461A1
; GENERAL INFORMATION:
; APPLICANT: DUKE UNIVERSITY
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS VACCINE
; FILE REFERENCE: 1579-548
; CURRENT APPLICATION NUMBER: US/09/775,805
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/497,497
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-775-805-44

Query Match 40.7%; Score 37; DB 9; Length 16;
Best Local Similarity 45.5%; Pred. No. 72;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 NRWEDPGKQLY 12
| |:: || :|
Db 5 NMWQEVGKAMY 15

RESULT 6
US-09-775-805-67
; Sequence 67, Application US/09775805
; Publication No. US20010036461A1
; GENERAL INFORMATION:
; APPLICANT: DUKE UNIVERSITY
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS VACCINE
; FILE REFERENCE: 1579-548
; CURRENT APPLICATION NUMBER: US/09/775,805
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/497,497

```
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-775-805-67

  Query Match             40.7%;  Score 37;  DB 9;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 72;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 7
US-09-775-805-75
; Sequence 75, Application US/09775805
; Publication No. US20010036461A1
; GENERAL INFORMATION:
;  APPLICANT: DUKE UNIVERSITY
;  TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS VACCINE
;  FILE REFERENCE: 1579-548
;  CURRENT APPLICATION NUMBER: US/09/775,805
;  CURRENT FILING DATE:  2001-02-05
;  PRIOR APPLICATION NUMBER: 09/497,497
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: HIV-1
;   OTHER INFORMATION: Th-dominant/subdominant CTL epitopes in MVA.
US-09-775-805-75

  Query Match             40.7%;  Score 37;  DB 9;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 72;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 8
US-09-775-805-89
; Sequence 89, Application US/09775805
; Publication No. US20010036461A1
; GENERAL INFORMATION:
;  APPLICANT: DUKE UNIVERSITY
;  TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS VACCINE
;  FILE REFERENCE: 1579-548
;  CURRENT APPLICATION NUMBER: US/09/775,805
;  CURRENT FILING DATE:  2001-02-05
;  PRIOR APPLICATION NUMBER: 09/497,497
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: HIV-1 Th-CTL
;   OTHER INFORMATION: A2 p17 epitope (A2 Variants) in MVA
US-09-775-805-89

  Query Match             40.7%;  Score 37;  DB 9;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 72;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 9
US-09-862-849-1
; Sequence 1, Application US/09862849
; Patent No. US20020013274A1
; GENERAL INFORMATION:
;  APPLICANT: Sudhir Paul
;  APPLICANT:  Larry J. Smith
;  APPLICANT:  Gennady Gololobov
;  TITLE OF INVENTION: Methods for Identifying Inducers and Inhibitors of Proteolytic
;  TITLE OF INVENTION:  Antibodies, Compositions and Their Uses
;  FILE REFERENCE: UNMC 63123 DIV
;  CURRENT APPLICATION NUMBER: US/09/862,849
;  CURRENT FILING DATE:  2001-08-29
;  PRIOR APPLICATION NUMBER: US 09/046,373
;  PRIOR FILING DATE: 1998-03-23
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Human Immunodeficiency Virus-1
US-09-862-849-1

  Query Match             40.7%;  Score 37;  DB 9;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 72;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 10
US-09-894-018-199
; Sequence 199, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
;  APPLICANT: EPIMMUNE, Inc.
;  APPLICANT:  Sette, Alessandro
;  APPLICANT:  Chestnut, Robert
;  APPLICANT:  Livingston, Brian
;  APPLICANT:  Baker, Denisw
;  APPLICANT:  Newman, Mark
;  APPLICANT:  Brown, David
;  TITLE OF INVENTION: METODS AND SYSTEM FOR OPTIMIZING
;  TITLE OF INVENTION:  MINIGENES AND PEPTIDES THEREBY
;  FILE REFERENCE: 39963-20033.00
;  CURRENT APPLICATION NUMBER: US/09/894,018
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: PCT/US00/35568
;  PRIOR FILING DATE: 2000-12-28
;  PRIOR APPLICATION NUMBER: US 60/173,390
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: US 60/284,221
;  PRIOR FILING DATE: 2001-04-16
;  NUMBER OF SEQ ID NOS: 368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 199
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Transgenic mouse
US-09-894-018-199

  Query Match             40.7%;  Score 37;  DB 9;  Length 16;
```

```
  Best Local Similarity   45.5%;  Pred. No. 72;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          6 NMWQEVGKAMY 16


RESULT 11
US-09-894-594-66
; Sequence 66, Application US/09894594
; Publication No. US20030017497A1
; GENERAL INFORMATION:
;  APPLICANT: Kieber-Emmons, Thomas
;  APPLICANT:  Weiner, David B.
;  APPLICANT:  Monzavi-Karbassi, Behjatolah
;  TITLE OF INVENTION: Peptide Mimotopes of Carbohydrate Antigens and DNA Molecules Encc
;  TITLE OF INVENTION:  Same
;  FILE REFERENCE: UPN-3984
;  CURRENT APPLICATION NUMBER: US/09/894,594
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: 09/601,558
;  PRIOR FILING DATE: 2000-11-07
;  PRIOR APPLICATION NUMBER: PCT/US99/02405
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 60/073,690
;  PRIOR FILING DATE: 1998-02-04
;  PRIOR APPLICATION NUMBER: 60/214,517
;  PRIOR FILING DATE: 2000-06-28
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 66
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Novel Sequence
US-09-894-594-66

  Query Match             40.7%;  Score 37;  DB 10;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 72;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 12
US-10-103-395-230
; Sequence 230, Application US/10103395
; Publication No. US20020160019A1
; GENERAL INFORMATION:
;  APPLICANT: EPIMMUNE, Inc.
;  APPLICANT:  Sette, Alessandro
;  APPLICANT:  Sidney, John
;  APPLICANT:  Southwood, Scott
;  TITLE OF INVENTION: IDENTIFICATION OF BROADLY REACTIVE DR
;  TITLE OF INVENTION:  RESTRICTED EPITOPES
;  FILE REFERENCE: 39963-20016.01
;  CURRENT APPLICATION NUMBER: US/10/103,395
;  CURRENT FILING DATE:  2003-01-03
;  PRIOR APPLICATION NUMBER: US 09/009,953
;  PRIOR FILING DATE: 1998-01-21
;  PRIOR APPLICATION NUMBER: PCT/US98/01373
;  PRIOR FILING DATE: 1998-01-23
;  PRIOR APPLICATION NUMBER: US 60/036,713
;  PRIOR FILING DATE: 1997-01-23
;  PRIOR APPLICATION NUMBER: US 60/037,432
;  PRIOR FILING DATE: 1997-02-07
;  NUMBER OF SEQ ID NOS: 274
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 230
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-103-395-230

  Query Match             40.7%;  Score 37;  DB 13;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 72;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          6 NMWQEVGKAMY 16


RESULT 13
US-10-114-716A-1
; Sequence 1, Application US/10114716A
; Publication No. US20030078203A1
; GENERAL INFORMATION:
;  APPLICANT: Sudhir Paul
;  APPLICANT:  Yasuhiro Nishiyama
;  TITLE OF INVENTION: Covalently Reactive Transition State
;  TITLE OF INVENTION:  Analogs and Methods of Use Thereof
;  FILE REFERENCE: UTH001HB
;  CURRENT APPLICATION NUMBER: US/10/114,716A
;  CURRENT FILING DATE:  2002-04-01
;  PRIOR APPLICATION NUMBER: 09/862,849
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: 09/046,373
;  PRIOR FILING DATE: 1998-03-23
;  PRIOR APPLICATION NUMBER: 60/280,624
;  PRIOR FILING DATE: 2001-03-31
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Human Immunodeficiency Virus-1
US-10-114-716A-1

  Query Match             40.7%;  Score 37;  DB 14;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 72;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 14
US-10-041-414-42
; Sequence 42, Application US/10041414
; Publication No. US20030087225A1
;   GENERAL INFORMATION:
;        APPLICANT: SHIVER, JOHN W.
;                   DAVIES, MARY ELLEN
;                   FREED, DANIEL C.
;                   LIU, MARGARET A.
;                   PERRY, HELEN C.
;        TITLE OF INVENTION: SYNTHETIC HIV ENV GENES
;        NUMBER OF SEQUENCES: 53
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
;             STREET: 126 E. LINCOLN AVE., - P.O. BOX 2000
;             CITY: RAHWAY
;             STATE: NEW JERSEY
;             COUNTRY: US
;             ZIP: 07065-0907
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/041,414
;             FILING DATE: 08-May-2002
;             CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/802,368
;             FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;             NAME: HAND, J. MARK
;             REGISTRATION NUMBER: 36,545
;             REFERENCE/DOCKET NUMBER: 19643
;         TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 732-594-3905
;             TELEFAX: 732-594-4720
;    INFORMATION FOR SEQ ID NO: 42:
;         SEQUENCE CHARACTERISTICS:
;             LENGTH: 16 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-041-414-42

  Query Match             40.7%;  Score 37;  DB 14;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 72;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          5 NMWQEVGKAMY 15


RESULT 15
US-10-371-525-308
; Sequence 308, Application US/10371525
; Publication No. US20030203869A1
; GENERAL INFORMATION:
;  APPLICANT: Fikes, John D.
;  APPLICANT:  Hermanson, Gary G.
;  APPLICANT:  Sette, Alessandro
;  APPLICANT:  Ishioka, Glenn Y.
;  APPLICANT:  Livingston, Brian
;  APPLICANT:  Chesnut, Robert W.
;  APPLICANT:  Epimmune Inc.
;  TITLE OF INVENTION: Expression Vectors for Stimulating an
;  TITLE OF INVENTION:  Immune Response and Methods of Using the Same
;  FILE REFERENCE: 39963-20022.01
;  CURRENT APPLICATION NUMBER: US/10/371,525
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: US 09/311,784
;  PRIOR FILING DATE: 1999-05-13
;  PRIOR APPLICATION NUMBER: US 60/085,751
;  PRIOR FILING DATE: 1998-05-15
;  NUMBER OF SEQ ID NOS: 463
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 308
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: HIV1 ENV 566 (peptide F091.15)
US-10-371-525-308

  Query Match             40.7%;  Score 37;  DB 15;  Length 16;
  Best Local Similarity   45.5%;  Pred. No. 72;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NRWEDPGKQLY 12
              | |:: || :|
Db          6 NMWQEVGKAMY 16
```

Search completed: August 25, 2005, 00:16:42
Job time : 158 secs

This Page Blank (usp[illegible]